

# Psychological Review

EDITED BY  
CARROLL C. PRATT  
PRINCETON UNIVERSITY

---

## CONTENTS

<i>Pierre Janet: 1859-1947:</i> E. R. GUTHRIE .....	65
<i>The Postulates and Methods of 'Behaviorism':</i> KENNETH W. SPENCE ...	67
<i>Factor Analysis in a Test-Development Program:</i> J. P. GUILFORD .....	79
<i>On a Distinction Between Hypothetical Constructs and Intervening Variables:</i> KENNETH MACCORQUODALE and PAUL E. MEEHL .....	95
<i>Reaction to Frustration—A Critique and Hypothesis:</i> S. STANSFELD SARGENT	108
<i>Relations Between Philosophy and Psychology:</i> ALBERT G. A. BALZ ....	115

---

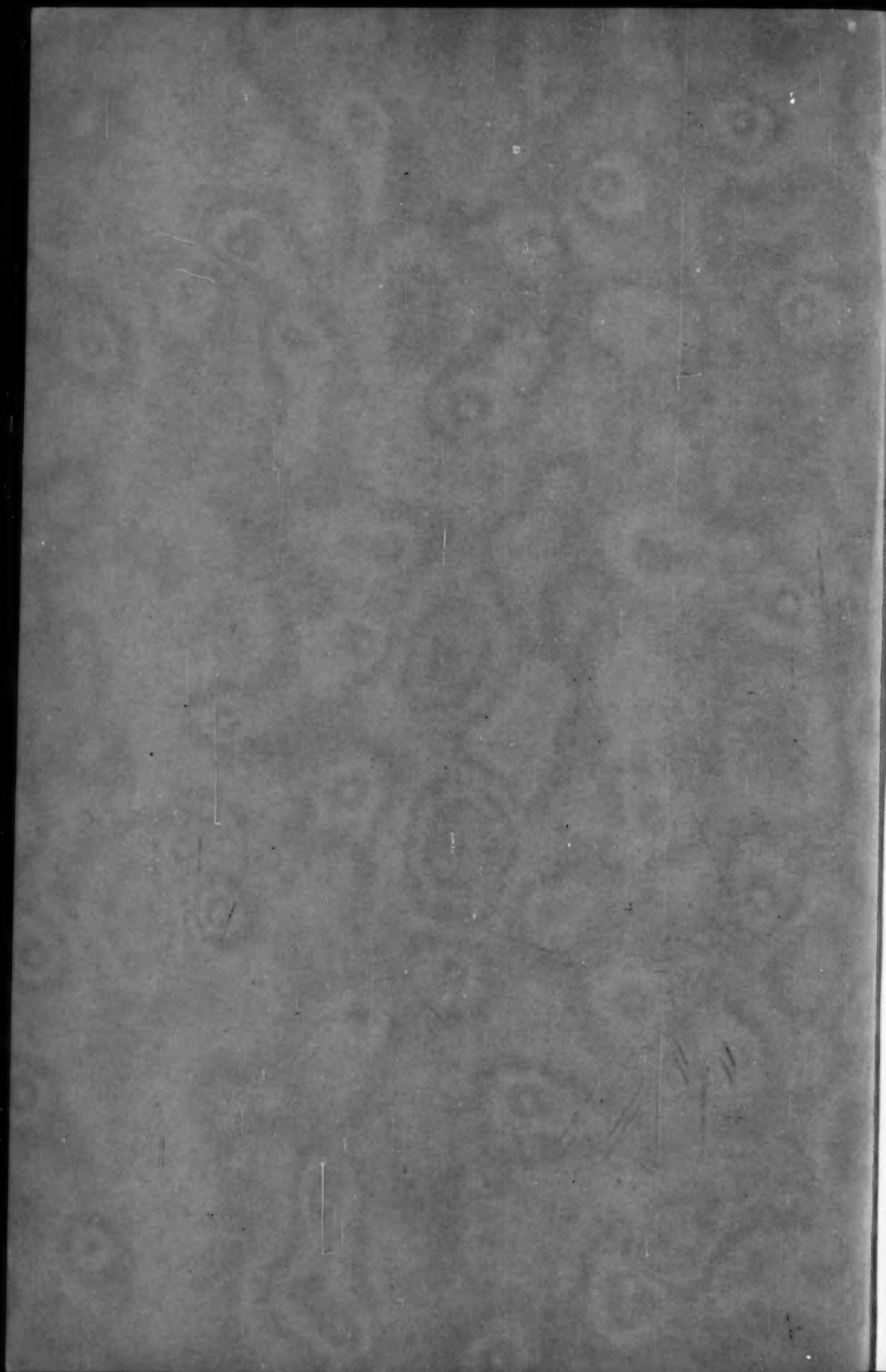
PUBLISHED BI-MONTHLY BY THE  
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.  
PRINCE AND LEMON STS., LANCASTER, PA.  
AND 1515 MASSACHUSETTS AVE., N. W., WASHINGTON 5, D. C.

\$5.50 volume

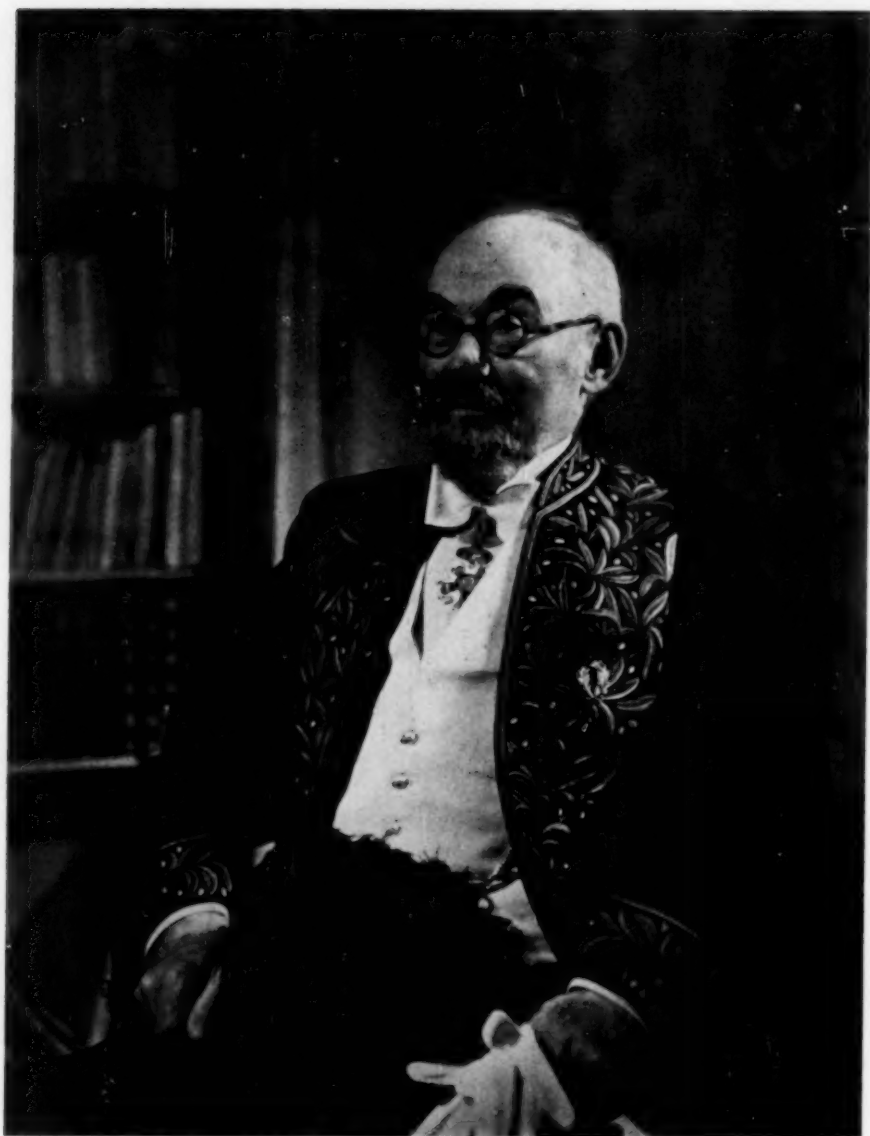
\$1.00 issue

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in the Act of February 28, 1935, embodied in paragraph 4, Section 538, P. L. and R., authorized Jan. 8, 1948







P. Janetz



# THE PSYCHOLOGICAL REVIEW

PIERRE JANET

1859-1947

This year saw the death of a man whose name will be one of the great landmarks in the history of psychology—Pierre Janet. Janet was born in Paris in 1859. He received his doctorate in letters in 1889 and in medicine in 1893. In 1890 he joined the Clinic at Salpêtrière under Charcot. During the remainder of his career he was identified with that Clinic and with the Collège de France and the University of Paris. He lectured in the Harvard Medical School in 1906. His last appearance in the United States was at the Harvard Tercentenary where he delivered one of the most spirited and interesting addresses of the occasion.

His remarkably varied and extensive publications began in 1886 with publication of a study on the psychology of Malebranche and the theory of animal mind in the seventeenth century. From then on his publications include in their titles all the important concepts of abnormal psychology—systematized anesthesia, aptomatisms, aboulia, fixed ideas, hysteria, obsessions, psychasthenia, possession and fugues, the nature of the personality, surgery during hypnosis, neurosis, the amnesias, déjà vue, and psychoanalysis. Among these there appeared in 1893 a small manual of philosophy. In 1915 he began to develop

his theory of psychological tension which marked an extremely promising new line of thought in the understanding of human behavior.

It was in his *Les Médications Psychologiques*, published in the early twenties and translated into English by Cedar and Eden Paul under the title, *Psychological Healing*, that Janet developed his general theory of the psychoneuroses. In these two volumes his general ideas are illustrated with the most astonishing wealth of detail from his own personal experiences at Salpêtrière. There runs through both theory and its illustrations a remarkable clarity and adherence to the general concepts of science. At no point in his theory has he appealed to powers of darkness like those invoked by Freud. Being himself probably more responsible than any other one psychologist for the modern use of the term "the unconscious," Janet has an expressed dismay at what he called its "*trop belle destinée*." He reminds us of something often forgotten—that in using the term "the unconscious," our only practical and scientific reference must be to a class of actions which can be seen or heard or otherwise directly observed.

His notion of "*force mentale*" in terms of which he explains the development of the psychoneuroses, and the efficacy of various methods of their cure, he does not reduce to its physiological basis; but he makes the assumption that

*Note:* The manuscripts for this and the January issue were accepted and prepared by the former editor, Herbert S. Langfeld, in the absence in Turkey of the present editor.

it has such a basis. A very typical illustration of the difference between his approach and the psychoanalytic approach lies in his observation that in his experience hysterics, whose behavior would be explained by Freud in terms of strong sexual tensions, are more properly described as persons lacking in sex drive or persons whose margin of energy is so small that they must learn to be self-centered and ungenerous. Probably the chief defect in Janet's psychological system is its failure to emphasize the dependence of personality on past experience, personality as the result of adaptation to past circumstances.

The astonishing range of Janet's thinking and insight can be appreciated only after an acquaintance with the volumes in which were published the lectures which he gave in a long term of years at the Collège de France. Two of these volumes, one of them on

*l'Évolution de la Mémoire et de la Notion du Temps*; the other, *l'Évolution Psychologique de la Personnalité*, published in 1928 and 1929 from stenographic notes, examine the psychological origins of many of our basic concepts in the fields of time and personality.

Many of the problems raised in his lectures are no longer occupying the attention of psychologists or philosophers, but many of them might well be occupying that attention. If Janet's name disappears from current mention in psychological writing as thoroughly as has Wilhelm Wundt's, the loss to psychology will probably be greater. The development of our science will have neglected certain promising lines of attack on the nature of disorders of the personality which Janet opened up but which now find themselves neglected.

E. R. GUTHRIE

University of Washington

# THE POSTULATES AND METHODS OF 'BEHAVIORISM'<sup>1</sup>

BY KENNETH W. SPENCE

*State University of Iowa*

There was a time when the term 'behaviorism' in the title of a speech required no further specification. Every psychologist at least knew the referent to be that new brand of psychology, introduced by Watson, which proposed to break with tradition and deny that psychology had anything to do either with a mentalistic entity called consciousness or a method known as introspection. Today the situation is not so simple. The term 'behaviorism' may, on the one hand, merely imply a very general point of view which has come to be accepted by almost all psychologists and thus does not point to any particular group or theoretical position. Or, on the other hand, it may refer to any one of several varieties of behaviorism which have been offered as supplementations or modifications of the original formulation of Watson (*e.g.*, molecular behaviorism, molar behaviorism, operational behaviorism, purposive behaviorism, logical behaviorism—to mention only some of the varieties). While these current formulations usually acknowledge some debt to Watson, for various reasons which we cannot stop to discuss they almost invariably take great pains to differentiate themselves from what has come to be known as 'Watsonian Behaviorism' or 'Watsonianism.' In fact, so far as I know, there are no proponents today of the original Watsonian version. Proper care should be taken to note, however,

that this statement holds true only for the particular pattern of assumptions that Watson advanced. Many of the basic postulates of his formulation are to be found in the present-day varieties of behaviorism and, what is more important, probably, in the underlying working assumptions of the great majority of present-day American psychologists.

Now that I have taken the precaution to differentiate the behaviorisms of today from the original version of behaviorism, I should like to call attention to the further interesting fact that with the exception possibly of Tolman very few, if any, current psychologists ever seem to think of themselves, or at least explicitly refer to themselves, as behaviorists. Such labeling, when it occurs, is usually the contribution of psychologists who consider themselves opposed to behaviorism. Undoubtedly, one of the reasons underlying this absence or lack of 'old-school-tie' spirit is that a large majority of present-day American psychologists just take for granted many of the behavioristic assumptions and, occupied as they have been with the details of developing and applying their specific research tools, they have had little time or inclination to give much thought to the more general methodological and systematic problems of their science.

Even the more theoretical-minded of the behavioristically-oriented psychologists seem to have been too preoccupied with matters of detail to get around to the consideration of a more general theoretical framework. Instead of attempting to formulate a complete system of psychology, these theorists have

<sup>1</sup> This article was an address given at the Symposium on 'The Postulates and Methods of Gestalt Psychology, Behaviorism and Psychoanalysis' given at the Conference on Methods in Philosophy and the Sciences in New York City, November, 1946. Some minor changes have been made in the paper itself and a list of references has been added.

been more concerned with the elaboration of relatively specific hypotheses concerning rather limited realms of data—e.g., theories of simple learning phenomena, motivational theories, theories of personality development, etc. As a consequence we find that instead of being built up around the symbol 'behaviorism,' allegiances tend to become attached to such labels as associationism, conditioning, reinforcement theory, frustration hypothesis, etc. It seems, in other words, that these psychologists have outgrown the stage of schools.

Under these circumstances, I cannot and I shall not undertake to present a fixed set of articles of faith, articulately and self-consciously held by a group of men calling themselves behaviorists. Instead, I shall attempt to formulate a few methodological principles that are, I believe, exemplified in the work of certain contemporary psychologists who would undoubtedly acknowledge a heavy historical debt to that earlier formulation known as the school of behaviorism.

The first problem that I shall discuss has to do with the behavior scientist's conception of the nature of psychological events. In the older, classical psychologies, whether of the structural or act varieties, the point of view taken was that psychology, if it was a natural science, was, to say the least, a somewhat unique one. Instead of being conceived like physics, for example, as concerning itself with events mediated by or occurring in the consciousness or immediate experience of the observing scientist, psychology was said to observe and analyze by a kind of inner sense immediate experience *per se*. Sensations, emotions, thoughts were regarded as observable aspects of direct experience rather than systematic constructs which, like the physicist's atoms and electrons, were inferred from immediate experience.

Fortunately, the relationship of im-

mediate experience (consciousness) to the data and constructs of science has been considerably clarified in recent years by the writings of several different groups of thinkers. The philosophers of science, particularly the logical positivists (1, 5, 6, 7), philosophically-minded scientists such as Bridgman (3) and, within psychology, such writers as Boring (2), Pratt (15), and Stevens (18) have succeeded, I believe, in making the point that the data of all sciences have the same origin—namely, the immediate experience of an observing person, the scientist himself. That is to say, immediate experience, the initial matrix out of which all sciences develop, is no longer considered a matter of concern for the scientist *qua* scientist. He simply takes it for granted and then proceeds to his task of describing the events occurring in it and discovering and formulating the nature of the relationships holding among them.

Boring stated this matter very clearly for psychologists in his book of some years ago, *The Physical Dimensions of Consciousness*. He wrote: "Thus the events of physics, as Wundt said, are mediate to experience, which stands in the background as the dator of scientific data, unrealizable as reality except inductively. In the same way psychology must deal with existential reals which are similarly mediate to experience. There is no way of getting at 'direct experience' because experience gives itself up to science indirectly, inferentially, by the experimental method" (2, p. 6).

More recently Pratt, in his *Logic of Modern Psychology* (15), has hammered home this same point with considerable effectiveness. As he points out, the subject matter of psychology is exactly the same in kind as all other sciences; any differentiation among the sciences is merely a matter of conven-



ience, a division of scientific labor resorted to as the amount of detailed knowledge increases beyond the capacity of a single person's grasp.

I think that it is of some historical interest to note in connection with this point that in the first of his articles introducing the behavioristic position, Watson took essentially the same stand. He wrote: "It [psychology] can dispense with consciousness in a psychological sense. The separate observation of 'states of consciousness' is, on this assumption, no more a part of the task of the psychologist than of the physicist. We might call this the return to a non-reflective and naive use of consciousness. In this sense consciousness may be said to be the instrument or tool with which all scientists work" (21, p. 176).

Acknowledging, then, that the psychologist conceives his task as that of bringing order and meaning into the realm of certain events provided by immediate experience, we now turn to the question of what these particular observed events are. In attempting to answer this question, attention should first be directed to the fact that the sense events in the experience of the observing scientist may depend upon or result from two different classes of conditions, intra-organic and extra-organic, the former exciting the interoceptors and the latter, the exteroceptors. The physical sciences, it should be noted, moreover, deal only with events of an extra-organic origin—i.e., those received through the exteroceptors. The data of classical psychology, on the other hand, were regarded as involving primarily sense events initiated through the interoceptors. These latter were regarded as being stimulated by such internal mental activities as thinking, desiring, emotional reactions, perceiving, etc., and hence were thought of as providing primary data concerning them.

It is apparent, however, that these internally initiated experiences differ rather markedly from the externally aroused ones in the extent to which they are publicly controllable and communicable. At least, if we can judge from the interminable disagreements of the introspective psychologists themselves, this class of experiences does not meet too well the requirements of social verification and acceptance demanded by the scientist. It was in the face of this difficulty that Watson made his suggestion that the psychologist, like all other scientists, should confine himself to those segments of his experience which have their origin in extra-organic conditions. In other words, the events studied by the psychologist, Watson held, should consist in observations of the overt behavior of *other* organisms, other persons than the observing scientist himself, and not in the observation of the scientist's own internal activities.

As everyone knows, however, most behavior scientists have continued more or less to make use of this latter type of material in the form of the objectively recordable verbal reports of their subjects. Indeed, the scientist himself, in certain circumstances, may assume a dual role and serve as both subject and experimenter. In this event his own introspective report is recorded as a linguistic response and becomes a part of the objective data. To some critics of the behavioristic viewpoint, this acceptance of the verbal reports of their subjects as a part of the data has seemed to represent an abandonment of the strict behavioristic position and a return to the conception that psychology studies *experiential* events as well as overt behavior.

Such a contention, it seems to me, fails to note a very important difference in the two positions. The introspectionist, it should be recalled, assumed a strict one-to-one relationship between

the verbal responses of his subjects and the inner mental processes. Accordingly, he accepted these introspective reports as *facts* or *data* about the inner mental events which they represented. The behavior scientist takes a very different position. He accepts verbal response as just one more form of behavior and he proposes to use this type of data in exactly the same manner as he does other types of behavior variables. Thus he attempts to discover laws relating verbal responses to environmental events of the past or present, and he seeks to find what relations they have to other types of response variables. He also makes use of them as a basis for making inferences as to certain hypothetical or theoretical constructs which he employs. In contrast, then, to the introspectionist's conception of these verbal reports as mirroring directly inner mental events, *i.e.*, facts, the behaviorist uses them either as data in their own right to be related to other data, or as a base from which to infer theoretical constructs which presumably represent internal or covert activities of their subjects. We shall return later to the use made of such language responses in the theorizing of the behaviorist.

From this all too cursory discussion of the initial data of the behavioristic psychologist, I should like now to turn to a consideration of the nature of the concepts which he employs to record and describe these events. I do not believe it is necessary for me to discuss at any length the position of the behaviorist with respect to the movement known as operationism. The insistence of the early behaviorists on a thoroughgoing operational analysis of the traditional mentalistic concepts was really nothing more than an anticipation of this somewhat overemphasized program. That a body of empirical knowledge cannot be built up without providing for verifi-

ability of the terms in use is simply taken for granted by the behaviorist. Instead, then, of talking about operational definition of psychological concepts, I should like to discuss certain matters related to a second criterion of acceptability of a scientific concept—namely, its *significance*.

One often hears criticisms to the effect that behavioristic concepts are too elementaristic, too atomistic, or that they fail to portray the real essence or true meaning of man's behavior. These latter critics often complain bitterly about the impoverishment of the mind, and of the lack of warmth and glowing particulars in the behaviorist's picture of psychological events. Some of these criticisms merely reflect, of course, a lack of appreciation on the part of some 'psychologists' as to the difference between scientific knowledge of an event on the one hand and everyday knowledge, or the kind of knowledge the novelist or poet portrays, on the other. Either by reason of training or because of their basically non-scientific interests, these critics have never really understood the abstract character of the scientific account of any phenomenon. The only reply that can be made to such a critic is to point out that the scientist's interests are quite different from his. There are, of course, other legitimate interpretations of nature and man than the scientific one and each has its right to be pursued. The behavior scientist merely asks that he be given the same opportunity to develop a scientific account of his phenomena that his colleagues in the physical and biological fields have had. If there are aspects of human or animal behavior for which such an account cannot ever be developed, there are not, so far as I know, any means of finding this out without a try. Unfortunately, the attitudes of too many psychologists with regard to this matter are not such as are

likely to lead them to the discovery of such knowledge. The difficulty, I fear, is that too many persons whose interests are non-scientific have become psychologists under the mistaken impression that psychology is one of the arts.

As to the criticisms that the behaviorist's concepts are too elementaristic, I must confess to the belief that the term 'elementarism' is merely one of those stereotypes, or 'rally-round-the-flag' words which the Gestalt psychologist has used in the defense and exposition of his holistic doctrines. However fervently the Gestalt psychologist may claim that he deals only with wholes, with total situations, the fact remains that if he is interested in discovering uniformities or scientific laws he must, of necessity, fractionate or abstract out certain features of the total events he observes. Such uniformities or laws describe ways in which events repeat themselves. Total concrete events, however, are seldom if ever repeated. Only certain features of events are repeated and since this is the case science must always abstract.

The problem here is really one of the size of the 'units of description' that the scientist is to employ and this brings us back to the criterion of acceptability of a scientific term which we referred to as *significance*. By the *significance* of a scientific concept is here meant the extent to which a concept or variable aids or enters into the formulation of laws. Significant concepts in science are those which are discovered to have functional relations with other concepts. Unfortunately, there are few if any rules for deciding *a priori* which concepts will and which ones will not be significant. Whether elementaristic concepts or units of description which, like the Gestaltists, are nearer the 'meaningful' common sense level, are to be chosen is entirely a pragmatic matter of which ones are most successful—i.e., which

ones lead to the discovery of laws. This can be ascertained only by trying them out.

Attention might also be called here to the further fact that it is entirely conceivable that different sizes or levels of descriptive units may be employed for the same set of events. The physical sciences provide us with numerous instances of this sort of thing and we see examples of it in psychology both in the description of behavior and stimulus events. Thus, employing the terms of Brunswik (4) and Heider (8), we may make use of either a proximal or distal account of the stimulus situation, and behavior may be described either in terms of movements (muscular patterns) or in terms of gross achievements. The particular alternative chosen, molecular or molar, depends upon the interest and purpose of the scientist, the kind of law he expects to find or use. As Hull (11) has pointed out in discussing this matter, some of the seeming disagreements among current psychologists are merely that one prefers to use more molar concepts than another.

Such different descriptions, however, do not necessarily represent fundamental disagreements. If the two systems of concepts should each be successful in leading to the discovery and formulation of laws, it should also be possible to discover coordinating definitions which will reveal the interrelations of the two systems. Or, as Hull (11) suggests, the postulates or primary assumptions of those working at a more molar level may ultimately appear as theorems in a more molecular description.

To sum up, then, the position which the behavior scientist takes with respect to the selection of the descriptive concepts to be employed in his science, recognizes (1) that the *significance* of a concept is to be measured in terms of the extent to which it leads to the formulation of laws about the phenomena;



(2) that a scientific law is always, in some greater or less degree, abstract in the sense that it refers only to certain properties of the events or sequence of events it describes and ignores other properties which are irrelevant to the particular momentary purpose; (3) that the method of elementary abstraction or analysis has been highly successful in all fields of science. While the disentanglement of the great complexes of properties and relations (sequences) among psychological events is undoubtedly much more difficult than in the case of physical phenomena, the difference between them need not be regarded as more than one of degree. On the basis of this assumption there would seem to be little reason for abandoning the method of abstraction or analysis.

We have said that the primary aim of the behavior scientist is to bring order and meaning into the particular realm of events he studies. Ordering a set of observable events for the scientist consists in discovering relationships between the events or, as we say, in the finding of empirical laws. The scientist seeks to establish laws relating his concepts or variables because they make possible explanation and prediction.

In the case of such areas of science as physics, the finding of empirical laws has involved chiefly the process of inductive generalization from observation and experimentation. In other words, in physics it has been possible to isolate sufficiently simple systems of observation to arrive at such laws in this manner. The situation in psychology and the other behavior sciences is quite different. Primarily because of the greater complexity of psychological as compared with physical phenomena, the psychologist has either been unable to isolate, experimentally, simple systems, or he has not found satisfactory means of measuring all of the relevant variables in the system under observation. In

this circumstance he has resorted to guesses or postulations as to the uncontrolled or as yet unmeasurable factors. As a result of this difference the term 'theory' has, as I have pointed out elsewhere (17), come to have a very different connotation in psychology from that which it has in physics. Theories in physics are constructions which serve primarily to integrate or organize into a single deductive system sets of empirical laws which previously were unrelated. The classical example is, of course, the Newtonian integration of the previously unconnected areas of mechanics and astronomy by the gravitational theory. Other well-known examples are the electro-magnetic theory of light and the kinetic theory of gases.

In psychology, on the other hand, theories serve primarily as a device to aid in the formulation of the empirical laws. They consist in guesses as to how the uncontrolled or unknown factors in the system under study are related to the experimentally-known variables. To these hypothetical constructs Tolman (20) has applied the very appropriate term 'intervening variable' because they are assumed to intervene between the measurable environmental and organic variables, on the one hand, and the measurable behavior properties on the other.

The manner in which the behavior scientist has used these hypothetical, intervening constructs may be shown by considering the various kinds of laws which the psychologist seeks to discover. Confining ourselves for the moment to laws which do not involve any hypothetical components, we find that the variables studied by the behavioristic psychologist fall into two, or possibly three, main groups:

- (1) Response variables: measurements of behavior properties.
- (2) Stimulus variables: measure-

ments of properties of the physical and social environment.

- (3) Organic variables: measurements of neuroanatomical or neurophysiological properties of the organism.

The different types of empirical relationships or laws in which psychologists have been interested are as follows:

1.  $R = f(R)$
2.  $R = f(S)$
3.  $R = f(O)$
4.  $O = f(S)$

Type 1 laws are laws of association of behavior properties. A great deal of use is made of the statistical constant, the coefficient of correlation, in the formulation of these laws and, as is well known, this type of law is investigated extensively in the field of psychological testing.

Type 2 laws may be concerned with the present environment or with past environmental events. Thus in the case of the typical perception experiments, we are interested in the effects of variation of aspects or features of the en-

vironmental stimulus on the perceptual or discrimination responses of the subject. Best examples of laws relating behavior to past events in the environment are laws of learning, laws of secondary motivation, etc.

For the most part the present-day behavioristic psychologists tend to concentrate their energies on these two classes of laws and to a very considerable extent they have favored the use of the molar rather than molecular concepts. A few psychologists whose interests have been in mediational problems have concerned themselves with type 3 and type 4 laws. These latter are obviously in the field of neurophysiological psychology and have in the main been concerned only with the simplest kinds of behavior phenomena—*e.g.*, sensory responses. Indeed, our inability to develop measures of this class of events (*i.e.*, organic variables) in the case of the more complex behavior phenomena has been one of the factors underlying the substitution of the hypothetical intervening constructs in their place.

Figure 1 continues this analysis of the laws of psychology. In this dia-

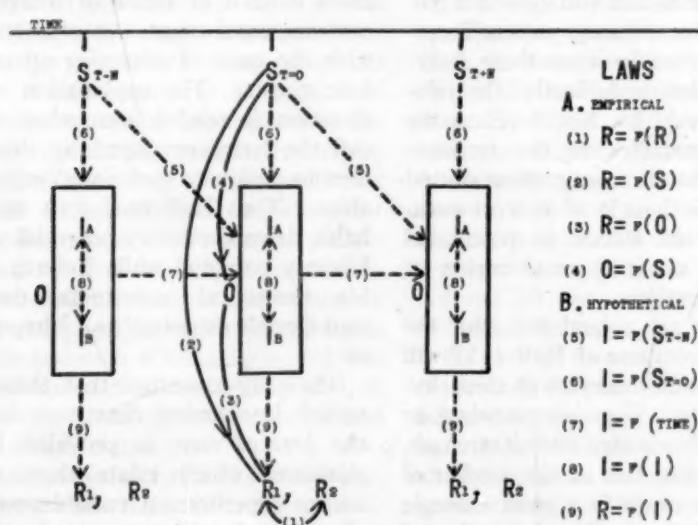


FIG. 1. Showing different kinds of laws.

gram I have attempted to portray, in addition to the four types of empirical laws which we have been discussing, the new hypothetical or guessed-at types of relationships which are involved in the introduction of the hypothetical intervening constructs. These latter are indicated as  $I_a$  and  $I_b$  and are represented as *hypothetical state variables* (enclosed within the rectangle). The environment or world situation at three different time intervals is represented by  $S_t - n$  (past)  $S_t = 0$  (present)  $S_t + n$  (future). These S's and also the R's represent empirical variables. I have also represented the class of experimental neurophysiological variables of the first figure by the symbol O, to the left of the rectangle. The four classes of empirical laws, listed at the right side of the figure, are represented by the solid curved lines. The guessed-at or postulated laws relating the hypothetical state variables ( $I_a$ ,  $I_b$ , etc.) to the various experimental variables are represented by the dotted lines. Thus No. 5 type of 'law' defines or introduces the intervening variables in terms of past events; No. 6 type relates them to the present environmental variables and No. 7 to time; No. 8 'laws' present interrelations assumed between these intervening variables, and, finally, the relations represented by No. 9 relate the intervening variables to the response variables. That is to say, these dotted lines should be thought of as representative of different classes of postulated relationships, not the usual notion of an S-R connection.

Those who are acquainted with the theoretical constructs of Hull (11) will recognize specific examples of these hypothetical laws. Thus his postulate or definition of the construct habit strength, or  $S^H R$ , as a function of the number of past reinforcements is a good example of class No. 5 'law.' His assumption of the nature of the manner in which H

and D interact to determine E falls in Class No. 8 and his postulate as to how the construct of reactive inhibition ( $I_R$ ) is assumed to change (disintegrate) with time is an instance of No. 7 type of 'law.' Incidentally, it will be noted that this last relationship is the only one which is similar to the so-called dynamic or process laws of physics. This type of law states or describes the laws governing the changes that occur within a system in time.

A question concerning these theoretical constructs that invariably seems to arise is whether they represent some kind of internal, presumably neurophysiological, process or state. The persistence with which misunderstanding arises on this point is truly surprising. It is probably to be explained in terms of the difficulty and resistance we have in shedding old, familiar meanings of words. In this connection it is not a little amusing to note that whereas Hull is usually accused of stuffing the organism with mythological brain states, Tolman, whose theoretical concepts have exactly the same formal structure as those of Hull—i.e., intervening variables defined in terms of independent environmental events—is often charged with the guilt of dreaming up mentalistic ghosts. The explanation of this situation is readily seen when we recall the terms employed by these two men to designate their intervening variables. Thus Hull used such words as habit, drive, excitatory potential and inhibitory potential while Tolman named his theoretical constructs, demands, sign-Gestalt-expectations, hypotheses, etc.

The only meanings that these theoretical intervening constructs have at the present time is provided by the equations which relate them to the known experimental variables—the environmental measurements on the one hand and the behavior measures on the

other. Such equations constitute the definitions of these terms.

The present role of these theoretical constructs we have said is to aid the psychologist in his search for the empirical laws relating behavior to the conditions determining it. In this sense they are a kind of calculational device which helps us to write the complete law describing the interrelations between all of the relevant experimental variables. In a recent article (17) on this problem of theory construction in contemporary psychology I called attention to the point that it is possible in the case of the theoretical formulation of simple learning behavior developed by Hull to substitute in the successive equations introducing the intervening theoretical constructs and obtain a single equation which states the response measure as a function of the several antecedent environmental variables. In this equation the intervening theoretical variables are represented among the parameters of the equation.

While both Tolman and I have emphasized the heuristic value of this type of theoretical construction in the formulation of the complete form of the laws, Hull (12) has called attention to another use which these constructs serve. Such constructs as habit and excitatory potential also provide, he claims, convenient, quantitative representations or indices of the particular complex of experimental variables for which they stand. Thus instead of having to state that the subject has had so many reinforcements in the situation under conditions in which the goal was of such-and-such a magnitude and was delayed for such-and-such a period, it is possible to substitute the calculated value of habit strength.

Finally, there remains the possibility, at least, that these intervening constructs may turn out to have their counterparts somewhere under the skin

of the organism. Hull in particular has been quite prone to accept this possibility and has not hesitated to add further statements about these constructs which suggest their possible locus and functioning in the nervous system. His justification, however, has always been that such conjectures provide experimental hints to persons interested in making such coordinations of our knowledge. His main theoretical efforts have been primarily at the molar-behavioral level.

In concluding this discussion of the theoretical framework of the behavioristic psychologist, I should like to emphasize that it is as yet only in a very primitive state of development, a fact which has unfortunately been lost sight of by many of the current critics of this position. The theorist in this field apparently has to choose between attempting to lay down the general theoretical framework of the whole range of behavior phenomena or working out the detailed nature of one small realm of data. Tolman has, for the most part, chosen the former alternative with the consequence that his treatment is characterized by an obvious lack of detailed specification of his theoretical constructs. Hull, on the other hand, has elected to follow the second method. His recent book, *Principles of Behavior*, dealt only with the most *simple* instances of laboratory learning phenomena, classical and instrumental conditioning, and he and his students are now engaged in extending the fundamental laws there discovered to the major phenomena of individual behavior.

So far as theoretical constructs are concerned, it is obvious that the simple behavior phenomena dealt with by Hull and other behavioristic-oriented psychologists have not required (to any great extent) a whole class of hypothetical intervening variables that must ultimately be postulated. Thus the



theoretical constructs in Hull's recent book—habit, excitatory and inhibitory potential, drive, etc.—are what might be referred to as *state variables*. Each of these constructs represents a hypothetical condition or state of the organism which is assumed to have resulted from and is defined in terms of the past interactions of the organism and its environment. In contrast the new theoretical constructs referred to above will represent, not states, but hypothetical, non-observable responses, implicit processes, occurring in the individual. Thus, in dealing with the more complex types of animal and human behavior, implicit emotional responses, covert verbal responses and not easily observable receptor-exposure and postural adjustments will have to be postulated in addition to these state variables. As yet only a bare beginning has been made in the use of such theoretical constructs—e.g., anxiety reactions and their secondary reinforcing effects (14), fractional anticipatory goal reactions as the basis of purposive behavior (9, 10).

It is in this realm of theorizing that the verbal reports of human subjects are likely to be of most use to the behavior theorist, for presumably these reports can be made the basis on which to postulate the occurrence of these inferred activities. There are, of course, many pitfalls in the use of such verbal reports and considerable caution needs to be exercised in their use. However, careful control and checking in terms of other, non-verbal responses should provide a means of detecting distortions, both deliberate and otherwise, in this source of data (16).

A discussion of behaviorism, especially when it occurs in conjunction with a symposium which includes Gestalt psychology, requires at least some comment on the distinction often made between field and non-field theories in psychology. The Gestalt psychologists,

in particular, have been very fond of this contrast and they have not hesitated to imply that their theoretical structures are similar in some respect to the type of field theory in physics represented by the Maxwell electromagnetic theory and Einstein's gravitational theory. In some instances the further implication has been made that behavioristic theories are a mechanical type of theory and as such are just as outmoded as the mechanistic theories of physics. Now I have often wondered what our theoretical brethren from the field of physics would think of these claims if perchance they were ever to take a serious look at these two groups of theories. Certainly the behavioristic theoretical structure I have been talking about uses neither the mechanical models—i.e., particles with their attracting forces—nor the type of mathematical equations that characterize a mechanical theory. Nor do I believe that there is anything even remotely resembling the field equations of Maxwell and Einstein in the theoretical formulations of the Gestalt psychologists. In the sense, then, in which the theoretical physicist understands the dichotomy, mechanical versus field theory, no such distinction, in my opinion, exists in psychology today.

If, on the other hand, the concept of field refers in psychology essentially to the notion of a system of interdependent variables, with its implication that the behavior of an organism at any moment is a resultant of the totality of relevant variables, then there is not to my knowledge any behavioristic theory today which would not also be a field theory. Furthermore, if we accept the additional notion that it is the pattern of interrelationships between the determining variables that is the crucial factor differentiating psychological field theories from non-field theories, I do not believe that the behavior theories

which I have been describing would fail to qualify as field theories. The hypothetical equations which Hull (11) postulates in the introduction of his theoretical constructs provide in precise mathematical form these very patterns of interrelationship. Finally, as to the characteristic of field theory emphasized by Lewin (13) under the principle of contemporaneity—namely, that the behavior at any moment is a function of the situation *at that moment only* and not a function of past or future situations,—I find it difficult to believe that any present-day psychologist believes that other conditions than those of the present moment determine the behavior of this moment. Even the psychoanalyst never held, as Lewin sometimes seems to imply, that past events somehow jump through time to determine the present behavior, but, instead, conceived of these past events leaving their effects in the organism and through them determining the behavior of the moment. The behaviorist takes exactly the same view of the matter.

The development of our science has not been helped, in my opinion, by such distinctions as field and non-field theory. A much more useful procedure would be to examine in detail these differing theoretical positions with a view to ascertaining to what extent they differ in the particular variables they believe to be relevant in a particular instance and what differences, if any, exist in their postulation as to the pattern of the interrelationships involved—*i.e.*, in the form of the hypothetical laws they assume. It is my personal belief that if this procedure were followed there would be much less in the way of specific disagreements to settle than is usually thought. I base this prediction not only on the well-known fact that the Gestaltists, psychoanalysts and behaviorists have to a considerable extent been interested in very different realms

of psychological phenomena and that hence their theories are not in competition with one another, but also on the fact that very little real theorizing, particularly in the matter of specifying the precise form of the interrelations between the variables, has actually been done. It is most imperative that psychologists attempt to formulate their theories in as precise and articulate a manner as possible, for it is only by means of such theorizing that psychology can hope, finally, to attain full-fledged scientific statehood.

## REFERENCES

1. BERGMANN, G. The subject matter of psychology. *Phil. Sci.*, 1940, 7, 415-433.
2. BORING, E. G. *The physical dimensions of consciousness*. New York: The Century Company, 1933.
3. BRIDGMAN, P. W. *The logic of modern physics*. New York: Macmillan Company, 1928.
4. BRUNSWIK, E. The conceptual focus of some psychological systems. *J. Unified Sci. (Erkenntnis)*, 1939, 8, 36-49.
5. CARNAP, R. Testability and meaning. *Phil. Sci.*, 1936, 3, 419-471; 1937, 4, 1-40.
6. —. *Philosophy and logical syntax*. London: Kegan Paul, Trench, Trubner & Co., Ltd., 1935.
7. FEIGL, H. Operationism and scientific method. *PSYCHOL. REV.*, 1945, 52, 243-246.
8. HEIDER, F. Environmental determinants in psychological theories. *PSYCHOL. REV.*, 1939, 46, 383-410.
9. HULL, C. L. Knowledge and purpose as habit mechanisms. *PSYCHOL. REV.*, 1930, 37, 511-525.
10. —. Goal attraction and directing ideas conceived as habit phenomena. *PSYCHOL. REV.*, 1931, 38, 487-506.
11. —. *Principles of behavior*. New York: D. Appleton-Century Co., 1943.
12. —. The problem of intervening variables in molar behavior theory. *PSYCHOL. REV.*, 1943, 50, 273-291.
13. LEWIN, K. Defining the "field" at a given time. *PSYCHOL. REV.*, 1943, 50, 292-310.
14. MOWRER, O. H. A stimulus-response

- analysis of anxiety and its role as a new forcing agent. *PSYCHOL. REV.*, 1939, 46, 553-565.
15. PRATT, C. C. *The logic of modern psychology*. New York: The Macmillan Company, 1939.
16. SKINNER, B. F. The operational analysis of psychological terms. *PSYCHOL. REV.*, 1945, 52, 270-278.
17. SPENCE, K. W. The nature of theory construction in contemporary psychology. *PSYCHOL. REV.*, 1944, 51, 47-68.
18. STEVENS, S. S. The operational definition of psychological concepts. *PSYCHOL. REV.*, 1935, 42, 517-527.
19. TOLMAN, E. C. *Purposive behavior in animals and men*. New York: The Century Company, 1932.
20. —. The determiners of behavior at a choice point. *PSYCHOL. REV.*, 1938, 45, 1-41.
21. WATSON, J. B. Psychology as the behaviorist views it. *PSYCHOL. REV.*, 1913, 20, 158-177.



# FACTOR ANALYSIS IN A TEST-DEVELOPMENT PROGRAM

BY J. P. GUILFORD

*University of Southern California*

The development of tests to meet the apparent requirements of selection and classification of personnel to fit particular assignments has traditionally proceeded along two lines. New tests have been (1) based upon the work-sample principle or (2) designed to measure hypothetical traits that are thought to be important as a result of a psychological job analysis. There is a third basis upon which test production and refinement can be accomplished, namely, the approach through factor analysis.

The aptitude-test development early in the Army Air Forces psychological program in World War II followed the traditional routes. The reasons were several. The great pressure for a valid classification battery at an early date favored the more direct and expedient steps and the exploitation of obvious leads. The factorial approach requires for its full exploitation an accumulation of intercorrelational information regarding tests and criteria. As this type of information grew, the basis for an increasing use of factor-analysis theory and practice was strengthened. There was, to be sure, much previous knowledge of factors, primarily from the work of Thurstone. As things turned out, it would have been very profitable to have utilized this as a starting point. The prior knowledge was by no means ignored, but the test-development program had been planned primarily from the traditional points of view.

In this article the emphasis will be upon the presentation of factor theory as a basis for rational test development and the advantages of factorial methods over other approaches. The dis-

cussion will be confined to one type of factor analysis—the Thurstone centroid method with rotation of reference axes—since experience has shown that that method gives the most psychologically meaningful results. Some illustrative findings from the AAF psychological program will be cited, with their implications. The discussion will begin with brief comments on the work-sample and the job-analysis modes of operation as a background for a presentation of the factorial approach. The three procedures can best be compared in answer to the question, "Why are tests valid?"

## I. WHY ARE TESTS VALID?

1. *Work-sample tests.* A work-sample test ordinarily presents a task that obviously resembles the features of a job or of some elementary component of the job. In the AAF classification battery, a good example of a work-sample test is known as the Complex Coordination Test. Developed between the two World Wars, it was apparently designed to present a task analogous to that of a pilot operating an airplane in flight.<sup>1</sup> The adjustment of a pilot's stick and rudder bar in response to changing signals confronting the examinee in this test is clearly similar in many respects to the activity of a pilot in controlling an airplane. Another test, known as Dial Reading, and still another, known as Table Reading, seem to duplicate in part the tasks of pilot, of navigator, and of bombardier, all of

<sup>1</sup> For full descriptions of tests mentioned in this article, see the AAF Reports (1).

whom must read dials and tables quickly and accurately as parts of their duties.

Work-sample tests rarely fail to exhibit some degree of validity. The extent of the validity will depend upon the completeness of the simulation of the job task and the relative importance of the task to the job as a whole. There is no question but that the most complete aptitude test would be the job itself. The best test of aptitude for learning to pilot an airplane would be to train the applicant in flying an airplane. A closely simulated task would be his performance in a good synthetic trainer for pilots. Both of these tests would be costly in terms of time, equipment, and, in the one case, possibly human lives. For these reasons, simpler, less expensive tests are to be recommended.

The use of a simpler work-sample test also presents some disadvantages. Even rather complex work-sample tests rarely cover the entire job for which selection is being made. A pilot, whether civilian or military, must be able to do more than maneuver an airplane safely. In order to cover all significant aspects of his job, other work samples must be brought into the picture. It may be supposed that if a sufficient number of work samples are included in a battery, covering all critical activities of any specialist, a fairly high degree of prediction would be attained. The number of activities in jobs, however, is rather large. The number of tests required to simulate them would be correspondingly large. The amount of duplication found in the complete coverage of a single job in this manner would probably be excessive and the practice would thus be wasteful. Even if only one work sample were required for each job, there would have to be as many tests as there are jobs or distinct types of jobs. That number is enormous.

While the validity of a work-sample test is usually taken for granted, there are sometimes puzzling results. No one would be amazed at the relatively high validity of the Complex Coordination test for selecting pilot trainees, with a typical validity coefficient of .39 (within the range of talent prevailing at the time aviation students were classified for aircrew training). It was designed as a pilot test, and the obvious similarity would seem to account for its predictive value in this connection. But it also proved to be equally valid for other assignments when the criteria had much less similarity to the test—a validity coefficient of .38 for air mechanics (when the criterion was a composite of academic grades and  $N$  was 300), and a validity coefficient of .40 for flexible gunners (when the criterion was a final academic examination and  $N$  was 173). It had lower, but substantial, correlations with accuracy of pistol firing (.30 when  $N$  was 350) and with carbine firing (.25 when  $N$  was also 350). Surely such versatility for a test needs to be explained. If it can be valid for predicting such disparate activities, ranging from academic grades to pistol firing, some explanation is demanded.

Similarly puzzling are the correlations of Complex Coordination with other tests which it does not at all resemble superficially. It was found to correlate almost as high with some paper-and-pencil tests (by name, Dial and Table Reading, .35, Instrument Comprehension, .36, and Mechanical Principles, .32) as it did with other psychomotor tests (.26 with Discrimination Reaction Time, .38 with Rotary Pursuit, and .32 with Rudder Control). Such facts leave one with a craving to understand these unexpected communalities of the Complex Coordination test. Such findings are not by any means confined to this test. It is my contention that only by making a study of the intercorrela-

tions of tests and criteria are we able to understand such mysteries. It is also contended that the understanding of communalities among tests leads to further enlightened progress in test development and in the discovery of new variables in human personality, as subsequent discussion will show.

2. *Tests based upon psychological job analysis.* A second answer to the question is that tests are valid because they measure hypothetical traits observed in making a job analysis. It requires some amount of job analysis, to be sure, to enable one to arrive at work-sample tests, but this hardly deserves the name 'job analysis' from the point of view of the psychological technologist. Psychological job analysis is designed to arrive at less superficial results. Its goal is to break a criterion activity down until the significant psychological abilities and traits have been identified. The abilities and traits so observed may be designated by common recognized psychological terminology.

In the analysis of the training of an aircraft pilot, such familiar categories as memory, attention, perception, judgment, comprehension, foresight, planning, coordination, fear, apprehension, and temperament were applied. Other traits that were used had little precedent in psychology, but, on the other hand, are more natural to the lay observer of aviation performance—"sense of sustentation, feel of controls, appropriateness of controls used," and the like. These categories could, no doubt, be further analyzed from the 'arm-chair' and could be translated into familiar psychological terms, such as tactual and kinesthetic perception.

Tests designed on the basis of this type of job analysis are variously successful. When they succeed, one has the basis for believing that the observation of the trait, its definition, and its translation into a test have all been correct.

If the tests turn out to be invalid for the purpose intended, the natural conclusion is that there was failure at one of the steps; which one is not immediately apparent. Factor analysis of such tests has shown, however, that a valid test constructed on this basis may actually be valid in spite of, rather than because of, the procedure followed. Some examples will support this statement. A test that was designed to measure supposed traits of speed of decision and reaction was found to measure factors identified as speed of perception, awareness of spatial relations, and psychomotor precision.<sup>2</sup> Further factor analysis may show that the supposed traits of speed of decision and of reaction may actually be represented by this test, but the evidence indicates that its entire validities for predicting the pilot, navigator, and bombardier criteria can be attributed to the three factors named. As another example, a test developed to measure practical judgment was found upon analysis to measure factors identified as verbal comprehension, mechanical experience, and reasoning, each to a higher degree than it measured a factor that could be called judgment. Tests developed to measure mechanical comprehension very often measure the power to visualize almost as much as they do a mechanical factor, and this mechanical factor is one of knowledge or experience rather than one of comprehension.

The prevailing approach to aptitude-test development to date has been closest to that of job analysis. Pre-war progress had gone very little beyond the procedures so well set forth by Hull about twenty years ago (4). Most of the success thus far achieved must be accredited to this mode of operation. From what is to follow, however, it will

<sup>2</sup> Definitions of factors mentioned in this article will be found in reference (1), Report No. 5, and more briefly in reference (3).

be seen that this approach has been largely a blind one and has progressed more by trial and error than by virtue of real understanding of what the requirements of prediction actually are and of what specific tests have to offer to meet those requirements.

3. *Tests based upon factor analysis.* The third answer to the question "Why are tests valid?" is that they measure in common with the job criterion certain fundamental factors. These variables must not be confused with the usual job-analysis categories, for they are derived in a very different manner. The remainder of this paper will be devoted to a brief account of the assumptions and theory underlying factor analysis as applied to test validity, and to the many advantages of this approach.

## II. RESUMÉ OF FACTOR THEORY

1. *Fundamental assumptions.* The first assumption of the factorial approach is that tests and criteria alike can be statistically analyzed into a limited number of basic traits that additively make up the total variance of each test or criterion. The term 'variance' may be regarded, roughly and simply, as merely a more exact expression for the idea of 'individual differences.' It should be emphasized that the analysis is statistical rather than observational in the ordinary sense. One reason why more of the factors have not been detected without the use of statistical procedures is that they are not obvious to surface inspection. For the most part, owing to the extreme complexity of the person observed and of his activities, they have eluded even the sophisticated observer. They are, however, usually recognizable and acceptable to most observers when they are pointed out after having been discovered by statistical analysis. The hiatus between a list of ordinary job-analysis traits and a list of statistically

derived factors is most striking. They coincide at some points, but are generally divergent.

To express the first assumption more exactly, the total variance of any test or of any criterion can be subdivided into the following component variances:

(1) a number of common factors (common in the sense that they appear in more than one test or criterion); (2) a possible specific factor (a factor that appears consistently in the same variable from time to time but not in other variables); and (3) error variance. Mathematically, if we let the total variance of a test or criterion be equal to 1.00, the same ideas can be expressed by the equation

$$x^2 = a_x^2 + b_x^2 + \dots + n_x^2 + s_x^2 + e_x^2 = 1.00 \quad (1)$$

in which

$x^2$  = the total variance of test X,

$a_x^2$  = the proportion of variance contributed by factor A,

$b_x^2$  = the proportion contributed by factor B,

$n_x^2$  = the proportion contributed by the  $n$ th factor,

$s_x^2$  = the proportion of specific variance, and

$e_x^2$  = the proportion that is error variance.

Thurstone has further defined the sum of the common-factor variance as the *communality*, which is described by the equation

$$h_x^2 = a_x^2 + b_x^2 + \dots + n_x^2. \quad (2)$$

The reliability of a test (its proportion of non-error variance) is described by the equation

$$r_{xx} = a_x^2 + b_x^2 + \dots + n_x^2 + s_x^2 = 1 - e_x^2. \quad (3)$$

It might be added here that what may appear to be specific variance in a test may prove to be one or more addi-



tional common factors when the test's relationships to new types of tests are taken into account. This statement does not preclude the possibility of genuine specific factors. In practice we have little interest in them, since nothing can be predicted from them outside the tests having them.

A second important assumption is that the intercorrelation of any two variables depends upon the factors that they have in common, and the factor weights or loadings. The general equation to describe this idea is

$$r_{jx} = a_j a_x + b_j b_x + c_j c_x + \dots + n_j n_x \quad (4)$$

in which

$r_{jx}$  is the correlation between test X and criterion J,

$a_j$  = the loading for factor A in criterion J,

$a_x$  = the loading for factor A in test X, and other symbols have analogous meanings.

A validity coefficient is therefore regarded as a summation of the cross products of the factor loadings of factors that test and criterion have in common. A factor loading is identical with the correlation of a measured variable with a factor.

From equation (4), several deductions can be drawn. It can be seen that the validity of a test or of a test battery can be maximized by taking several steps. One of these is to make sure that every factor that has a weight in the criterion also has a weight of like algebraic sign in the test or battery. Increasing the present validity of a test battery is best assured by adding coverage of an additional factor, or factors, not now represented. Another step toward maximal validity would be to increase the loading of a factor already measured in either test or criterion. Since a criterion is usually a relatively

fixed quantity, except as there are improved measures of it, this mode of improvement would have to come from improved tests. There are limitations to improvement from this direction, for maximal validity also requires optimal weighting of the various factors in keeping with their loadings in the criterion.

2. *Some general illustrations.* Let us apply the two fundamental assumptions just stated to some general illustrative situations of tests and a criterion. In Fig. 1 are shown in graphic form the

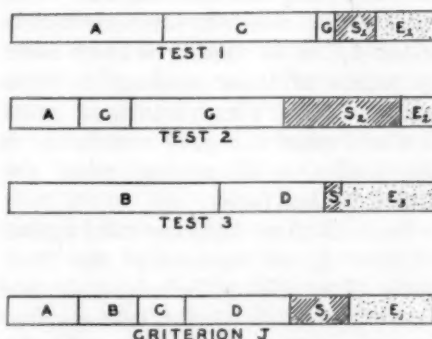


FIG. 1. Diagrams of the segregation of variances of tests and a criterion into component variances. Letters A to G stand for variances in common factors. S stands for a specific variance, and E for error variance, with subscripts of each in keeping with the total variable.

component variances of three tests and of a criterion J. The criterion has variances in four independent common factors, A, B, C, and D. Test 1 has variances in factors A and C; test 2 in factors A, C, and G; and test 3 in factors B and D. The proportions of these variances are indicated by sizes of areas of rectangles. Their composition of a total, along with specific and error variances, illustrates the summative equation previously given (equation 1).

It can be seen that all three tests have something in common with the criterion and may be regarded as being valid for the prediction of this criterion;

tests 1 and 2 by reason of sharing factors A and C, and test 3 by reason of sharing factors B and D. The degrees of validity can be readily estimated from equation (4), as follows:

$$\begin{aligned} r_{1j} &= a_j a_1 + c_j c_1 = (.4)(.6) \\ &\quad + (.4)(.6) = .48, \\ r_{2j} &= a_j a_2 + c_j c_2 = (.4)(.4) \\ &\quad + (.4)(.35) = .30, \\ r_{3j} &= b_j b_3 + d_j d_3 = (.3)(.7) \\ &\quad + (.5)(.5) = .46. \end{aligned}$$

Tests 1 and 3 are about equally valid, but for totally different reasons. Test 2 is valid by reason of the same factors as test 1, but its validity is much lower by reason of lower loadings in those factors. Test 2's high loading in factor G would make it a good contributor to the prediction of any criterion also loaded in that factor.

Since all three tests are valid against criterion J, let us consider the feasibility of combining them by pairs and note the relation of their common factors to multiple prediction. In this connection we must consider the intercorrelations of the tests. These can also be estimated from the common-factor loadings by using equation (4). They are:  $r_{12} = .45$ ,  $r_{13} = .00$ , and  $r_{23} = .00$ . Knowledge of the traditional multiple-regression equation would lead us to expect that the combining of tests 1 and 2 would bring about little improvement over the use of test 1 alone. It would also lead us to expect that the combination of tests 1 and 3 would lead to a very material improvement in prediction of the criterion over that of the best single test, namely, test 1. These expectations are fully born out by the results, and by reference to the diagrams the reasons will perhaps be more apparent. The multiple correlations are:  $R_{j,12} = .49$ , which is just .01 higher than the simple correlation  $r_{j1}$ ;  $R_{j,13} = .66$ , which is .18 higher than the same zero-order correlation,  $r_{j1}$ ; and  $R_{j,23} =$

.55, which is some improvement over the simple correlation  $r_{j3}$ . Here the reasons for desiring low correlations between tests when they both correlate positively with the same criterion which they are combined to predict are very clear.

As a matter of secondary interest here, if the variance in criterion J that is represented by  $s_j^2$  turns out to be another common factor, it would be desirable to determine its nature and to develop a good test of it that could be added to the battery. Let us say that  $s_j^2$  is equal to .14, or 14 percent of the total variance of the criterion. Then the loading corresponding to it, if the variance belongs to a single factor, would be the square root of .14, or .37. It would be well worth while to find a test, which, even if loaded only to the same extent with the factor, would add .14 to  $r_{j2}$ , the test validity.

The illustrations of factorial principles just given are in terms of simple, fictitious variables. Similar examples will be presented later in connection with actual AAF tests and criteria.

3. *The general nature of factors.* Before listing the advantages of developing tests on the basis of factorial knowledge, it is perhaps necessary to say something regarding the perennial issue as to the status of factors. Are they primary abilities, mathematical artifacts, real variables in personality, or culturally determined unities in behavior?

Operationally, factors are discovered from the systematic manner in which measures of individual differences intercorrelate. There seems no room for doubt as to the mathematical order so revealed, though there may be questions about specific findings in specific studies. Any finding is subject to further verification. Having found clusters of measures of individual differences that have much more in common with each

other than they have with other measures, the psychologist is usually not satisfied until he finds underlying 'reason,' or until he aligns the finding with other concepts, usually verbal; he tries to name the factor. The naming depends upon features that the clustering measures seem to have in common and that are unique to them. Attaching a meaningful label facilitates communication and systematic thinking. Any label or definition of a factor should be regarded as a hypothesis, in the same manner that any trait name is a hypothesis.

The least that can be said for factors is that they are convenient and dependable reference variables, derived by known operations that can be duplicated. This would seem to meet the best scientific requirements for research procedure. The dependability of factors has been amply demonstrated. The same factor will appear with comparable loadings in the same tests time after time. The factor known as perceptual speed was found to have loadings in a test entitled Speed of Identification (matching airplanes) as follows: .64, .58, .66, .67, .62, .65, .69, and .65. Analyses were based upon different groups of aviation students and involved different test batteries. A factor called spatial relations had loadings in the Complex Coordination test of .56, .50, .47, .47, .52, .52, .50, .46, and .46.<sup>3</sup> A general-reasoning factor appeared in Arithmetic Reasoning tests with loadings of .47, .47, .56, .50, .48, .47, .68, and .56. A factor denoted as visualization had loadings in a test called Me-

chanical Principles of .52, .49, .50, .39, .50, and .54. The verbal factor appeared in Reading Comprehension tests with loadings of .68, .63, .65, .65, .68, .69, and .73. The perceptual-speed factor was found correlated with the Complex Coordination test to the extent of .22, .22, .19, .20, .26, .22, .22, .28, .19, .25, and .14 in different analyses. The perceptual-speed loadings in the Mechanical Principles test were .00, .01, .03, .13, .12, .12, and .01 in as many analyses.

At this point it must be admitted that the criteria for rotation of reference axes in factor analysis are not completely objective. There is much room for the judgment of the investigator to play a part. Indeed, experience has shown that completely 'blind' rotations are not desirable. Use must sometimes be made of previous findings. But there are limits imposed upon subjective judgment by the observance of the goals of positive manifold and simple structure. Within these limits one cannot take extreme liberties. If more tests were factorially pure, those objective requirements alone would probably suffice. Some of the consistency reported above may thus be attributed to the use of cues from previous analyses. It is difficult for one who is experienced in factorial procedures, however, to believe that those consistencies could have been achieved without permission of the configuration of the factor structure itself. It would be a severe challenge to anyone to produce a very different set of consistent results if he worked within the limits of simple structure and positive manifold.

Factors, then, seem to be real dimensions of human personality when personality is defined as a phenomenon of individual differences. Factors are discovered by a set of mathematical operations from objective data, plus other operations that can be prescribed. They

<sup>3</sup> This spatial-relations factor is not identical with the space factor of Thurstone. AAF results show that his S factor is probably a composite of spatial relations and visualization. There are times when without enough definitive tests for a factor in a battery that is analyzed two factors may refuse to separate. This type of failure is not unique to factor analysis.



are reproducible by the same operations from data derived from new samples and even under somewhat varied conditions. They can usually be associated with verbal symbols which have psychological significance. They are 'primary' only in the sense of being reference variables which have a high degree of mutual independence. They probably correspond to observable variables or facts of biological or social origin or of joint bio-social origin. Which of these genetic cases applies will have to be decided for each factor. Some AAF results tend to show that genuine experience factors can be isolated, *e.g.*, a mechanical-experience factor and a mathematical-background factor. The status of other factors is still an open question.

### III. ADVANTAGES OF THE FACTORIAL APPROACH

1. *Precision.* The factorial approach provides an exact, quantitative picture of tests and criteria in terms of stable categories. The precision feature has already been suggested by the equations and by Fig. 1. More concrete illustrations are given in Figs. 2 and 3. Figure 2 shows the proportions of the factor variances in three of the AAF classification tests—the Reading Comprehension, Dial and Table Reading, and Discrimination Reaction Time tests. The first two of these are printed tests and the third is a psychomotor test. Although the Reading Comprehension test is primarily verbal, the verbal-comprehension factor takes up less than half of its total variance. Other factors that contribute materially are mechanical experience, two reasoning factors, and visualization. The presence of the mechanical variable here is understandable, though it could not have been decided with any assurance from inspection of the test. The reading selections were on technical material. Right an-

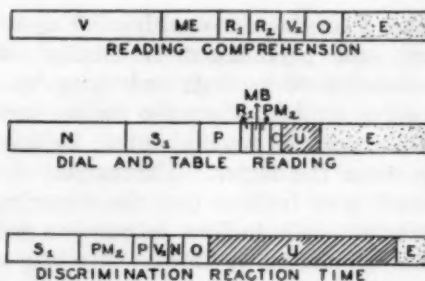


FIG. 2. Diagrams of the component variances of three Army Air Forces classification tests. The letters stand for:

- V—verbal-comprehension factor
- ME—mechanical-experience factor
- R<sub>1</sub>—reasoning I (general-reasoning) factor
- R<sub>2</sub>—reasoning II (common to analogies tests) factor
- V<sub>2</sub>—visualization factor
- O—other common factors, each with variance too small to mention separately
- U—unknown common-factor or specific-factor variances
- E—error variances
- N—numerical factor
- S<sub>1</sub>—space I (spatial-relations) factor
- P—perceptual-speed factor
- MB—mathematical-background factor
- M<sub>2</sub>—memory II (visual-memory) factor
- PM<sub>2</sub>—psychomotor II (precision) factor

swers to some items could be facilitated by the fact that the examinee had had acquaintance with mechanical concepts. The test had been constructed with the aim of requiring inferences from what

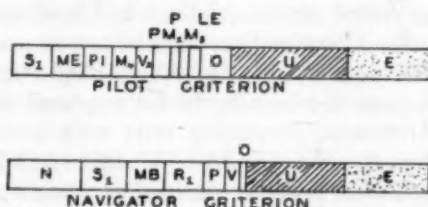


FIG. 3. Diagrams of the component variances of pilot and navigator training criteria. Letter symbols are as defined with Fig. 2, except for some additional ones:

- PI—pilot-interest factor
- M<sub>4</sub>—memory IV (content-memory) factor
- M<sub>3</sub>—memory III (picture-symbol association) factor
- PM<sub>1</sub>—psychomotor I (coordination) factor
- LE—length-estimation factor

was read, hence the reasoning variances. There was no attempt to measure visualization with this test, but there happened to be descriptive material in which comprehension was presumably facilitated by virtue of good visualizing ability. The small variances in other identifiable factors were less than two percent each. The sum of all common-factor variances fully reached the test's reliability coefficient, so there was no specific or unknown non-error variance.

Dial and Table Reading was designed as a work-sample test.<sup>4</sup> It is very complex factorially, which increases its chances of being valid but which makes its total scores ambiguous as to meaning. When an individual has a high score in it, we do not know whether he is particularly good in number ability, spatial ability, or perceptual-speed ability, or any combination of these. Only a small proportion of this test's non-error variance is still to be accounted for, as indicated by the portion labeled "U" in Fig 2.

The Discrimination Reaction Time test was designed to measure the job-analysis trait or traits of speed of decision and reaction. Its leading variance proved to be in spatial relations, followed by variance in psychomotor precision, perceptual speed and visualization. With a reliability of .92, this test has considerable unknown non-error variance to be accounted for.<sup>5</sup> Some of this unknown variance may,

indeed, be identifiable as speed of decision or speed of reaction. Whatever it is, however, it is not needed to account for this test's obtained validities for aircrew predictions.

Figure 3 shows two criteria analyzed into contributing variances of different kinds: common-factor, unknown, and error. The pilot criterion became much better known than the navigator criterion because many more experimental tests had been validated against it. There was opportunity to estimate the contribution of variances of some twenty-seven common factors to the pilot criterion, whereas for the navigator criterion there was opportunity to account for only 11 factors. It is estimated that some 52 per cent of the variance of the pilot pass-fail training criterion can be accounted for by contributions from twenty-three orthogonal factors. If we accept the estimate of reliability of the pilot criterion as being .80 (it is probably less than that) we can see that perhaps an addition of 28 per cent of the total variance could be predicted if we had the proper tests to measure the as yet unknown factors. Nine known factors account for about 56 per cent of the navigator criterion. How much the additional known common factors would increase this figure is a question still to be answered. It is safe to say that more of the non-error variance in the navigator criterion has been accounted for than in the pilot criterion. A comparison of the two criteria shows that they have very little in common; only the spatial-relations and perceptual-speed factors are material contributors to both criteria. This is a circumstance that is very favorable for differential selection, that is, classification.

2. *Economy.* The economy of the factorial approach lies in the relatively small number of variables with which one must deal in covering predictions.

<sup>4</sup> It is interesting to note that this test was developed first as two separate tests, but they were later combined when it was found that they functioned very similarly in predicting criteria. Factor analysis revealed that they had almost identical functional content.

<sup>5</sup> Since this test is an outgrowth of a very old laboratory technique, the finding that it is quite complex factorially should be of special interest to the experimental psychologist. No longer can the latter feel secure as to the nature of his experimental variables. Ambiguity of meaning of measurements is a much greater problem than has been generally recognized.

The number of distinct factors seems to be much larger than many have anticipated. In spite of this, however, by the use of factors the number of working variables is materially reduced, as compared with the number of tests generally employed. During the war, the AAF classification battery was composed of about twenty tests yielding as many scores, each of which received a weight (including some zero weights) in deriving a composite aptitude score for each AAF aircrew specialty. These twenty scores covered only about eight of the factors of the pilot criterion. This fact is eloquent of the wastage from overlapping of coverage of necessary variables. At the same rate, it would have required about sixty tests in order to cover the known factors of the pilot criterion. With a battery of pure tests, one for each factor, the number of tests can be reduced to the number of factors. Had aviation psychologists known what all the factors were at the beginning, their efforts could have been directed to the construction of far fewer than the one to two hundred tests that were actually designed and constructed. In future test-development programs, the AAF factorial findings should provide considerable guidance toward an economical continuation.

3. *Understanding of test validity.* The factorial approach enables us to understand why each test is valid for the prediction of some criteria and why it is not valid for the prediction of others. Figure 4 shows how the validity coefficients of each of two tests for the selection of pilots and navigators can be made intelligible. Equation (4) was applied to the factor loadings that test and criterion had in common in each case. The pilot validity of the Reading Comprehension test, it can be seen in Fig. 4, is almost entirely accounted for by secondary factors in the

test—mechanical experience, visualization, and spatial relations. The navigator validity of the same test is much greater and is accounted for mainly by other factors—verbal, reasoning, and number. In both instances just mentioned, the obtained validity is less than .01 greater than that expected from the factors and their loadings.

The pilot validity of the Dial and Table Reading test is mainly attributable to its space and perceptual-speed variances. The predicted validity is less than .02 short of the obtained validity. The navigator validity for this test is much greater, and the additional communality is to be found in the number, reasoning, and mathematical-background factors. In this instance, the predicted validity is nearly .02 greater than the obtained validity. All of these discrepancies are probably well within the limits of sampling errors. In a study of the closeness of obtained pilot validities to predicted validities, involving 90 tests, the median discrepancy was about .02. The correlation between predicted and obtained validity coefficients was .81. These results indicate merely the goodness of fit of data to the factor loadings estimated in tests and criterion. Loadings in the criterion had been estimated largely from validity coefficients; loadings in the tests from intercorrelations of tests.

4. *Construction of univocal tests.* A univocal test has its non-error variance confined to one common factor. In other words, it is factorially pure. The advantages of having univocal tests are many, particularly if the factors are themselves unrelated.<sup>6</sup> It should be apparent from what has been said thus far that factorially impure tests often

<sup>6</sup> In this discussion it will be assumed that the factors are orthogonal, *i.e.*, uncorrelated. There were some minor intercorrelations among factors, but not enough to detract from the arguments here presented.

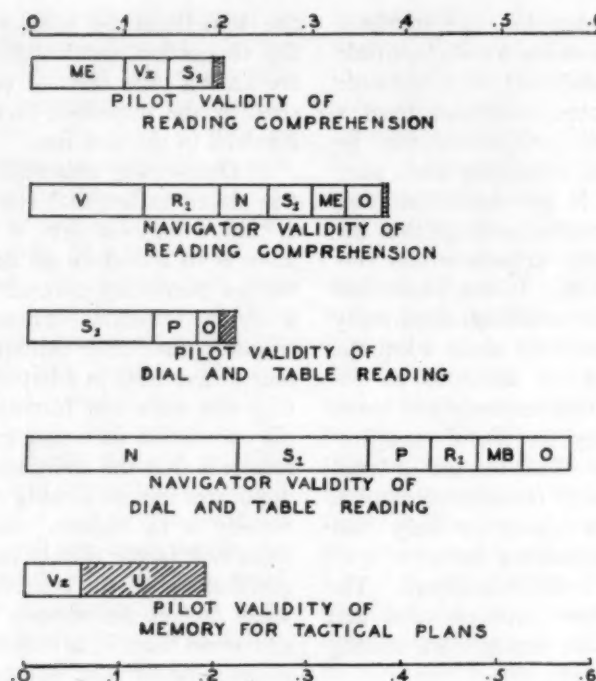


FIG. 4. Diagrams of factorial contributions to the validity coefficients of three tests for predicting pilot and navigator training criteria. Linear validity scale is shown at the top and bottom of the series of bars. The cross-hatched portions represent validity unaccounted for by known factors. Letter symbols have the same meanings as in Figs. 2 and 3.

contain variances that are unrelated to the criterion. The verbal and reasoning variances in the Reading Comprehension test and the number and reasoning variances in the Dial and Table Reading test add nothing to those tests for pilot selection. There is evidence that the verbal variance is even negatively correlated with the pilot criterion for the AAF aircrew training population. Vocabulary tests, which are most univocal for the verbal factor, consistently correlated negatively with the pilot criterion. An invalid variance in any test will lower the amount of correlation of the test with the criterion. It attenuates a correlation to the same extent as would a like amount of error variance. A factor with negative validity will

lower the correlation still more. From this point of view, it would seem important to rid a test of all common-factor variance except that in one factor.

Classification of personnel—and vocational guidance comes under this general heading of classification—requires univocal tests for most effective results. Differential predictions among general vocational choices which themselves possess different degrees of mutual independence is an exacting process. When vocational outlines are blurred and overlapping and when test scores by which we hope to sort individuals into vocational categories also overlap to a large extent, we are bound to make unsatisfactory decisions.

If we set up a single composite apti-



tude score for each job, the maximal degree of independence among aptitude scores is demanded. If each aptitude score is a weighted composite from a battery, maximal uniqueness can be achieved for each composite when pure tests are used. It was estimated from known common-factor loadings that the pilot and navigator criteria would correlate less than .20. It was impossible to determine this correlation empirically for the conditions were never adequate. A correlation of .20 indicates an extreme degree of independence and hence much opportunity for useful classification of trainees. With the use of tests, most of which were factorially complex, however, it was highly unlikely that this much independence between aptitude composites could be achieved. The actual correlations between pilot and navigator aptitude scores were usually of the order of .50.

Even a few pure tests in a battery will help a great deal, if they are properly chosen. For the use of single test scores in profiles, as in clinical vocational guidance, however, the ideal would be to have all tests univocal. Interpretations would then be much clearer and individual profile patterns would more nearly approach uniqueness.

The production of uni-factor tests is facilitated by the psychologist's awareness of the nature of factors. As experience with the factors increases, their contours become sharper and their features better defined. The purification of a test is a double process; of maximizing desired variance and minimizing undesired variance. Certain factors, such as verbal, numerical, and perceptual speed, are among the chief stowaways. It becomes relatively easy to tell beforehand when these factors are most likely to creep into a new test, and steps necessary to keep them out can be taken. Item-analysis procedures

can sometimes be used effectively in this connection, particularly when there are known pure tests of either the desired or the undesired factors that are involved in the new test.

5. *Discovering new valid factors.* It was stated earlier that one of the ways of improving validity of an aptitude score is to introduce an additional factor not previously covered in predicting a certain criterion. There has been a tradition that after combining the best four or five tests in a battery it is rarely that one more test increases the multiple correlation with that criterion. The reason is that the additional test merely duplicates factors already covered. The remedy is to discover and to measure some new factor that is valid—one that correlates with the criterion. Knowing what factors are already covered, one can avoid them in new tests and one has a much better idea whether a new test is likely to contribute something really new.

There are two ways in which one is led to new factors. One of these is to find that there is much unknown non-error variance in a test. This variance might be specific to the test, but probably is not. Such a finding is always a challenge to the factorist to see whether or not this variance is common to other tests or is common to some criterion. Excluding the features of the test that can be attributed to already better known factors, what does this test with unknown variance have that is unique? Several hypotheses are called to mind. New tests are developed, one or more with the attempt to bring out or to stress each supposed new factor. Analysis will show in which of these directions the greatest new communality lies. An example of this in the AAF was the discovery of the spatial-relations factor in the Complex Coordination test. This test was known to correlate with certain printed tests, which suggested that

it had a substantial amount of some intellectual variance. The leading hypotheses were (1) that this was mostly a matter of awareness of spatial arrangements and (2) that it was an ability to integrate a multitude of sensory impressions so as to yield a coordinated movement response. New printed tests were developed to fit the two hypotheses; space tests and integration tests. After analyses of these and other tests, it was decided that the first hypothesis was correct. The important practical outcome was not that a new valid factor was added to the list, but that a factor already included in the battery was better identified and it was found that this factor could be measured as well or even better by means of printed tests, with a great saving of time and equipment as compared to the psychomotor test that also measured it. Another outcome, however, was the discovery of three new factors in the integration tests, two of which are probably valid for pilot selection.

A second indication of the presence of a new factor is that the validity of a test is not fully accounted for by known factors. This was true of the Mechanical Principles test. Of all the mechanical tests developed, this one, which is similar to the Bennett test (2), had greatest pilot validity. Its validity was materially higher than that of a Mechanical Information test which was purest for the only factor that is unique to mechanical tests—the factor identified as mechanical experience. In preliminary analyses the additional valid factor in Mechanical Principles became recognized as visualization. With new tests constructed with the intent to measure visualization brought into the picture, the visualization hypothesis gained ground and was never contradicted by later results. The test called Memory for Tactical Plans, which is

represented in Fig. 4, was found to have a pilot validity of .19, of which the visualization factor can account for only .06. If a single factor accounts for the remaining .13 of the total validity, it must have not only a relatively high variance in the pilot criterion but also in the test. Because the test was developed as a memory test, stressing memory for verbal content (instructions for a mock mission), it is tempting to adopt this as a hypothesis. It was the only memory test of its kind developed and analyzed. It did not correlate materially with other memory tests, so this factor is either a radically different kind of memory or it is a non-memory factor. The content-memory hypothesis would be the one to follow up. In Fig. 3, this factor has been tentatively identified as M4 (Memory IV, since three others were identified as memory factors) and it has been given an estimated variance of .06.<sup>7</sup>

6. *Objective job analysis.* Ordinary job analysis, by direct observation or otherwise, taxes the best powers of the experienced psychologist. It may be the traditional psychological categories that are at fault. It may be a 'halo' effect, which applies to the psychological evaluation of jobs as well as to the rating of persons. However this may be, the list of factors does not coincide very well with the usual list of job-analysis categories.

A comparison of the two approaches to job analysis is decidedly in favor of the type represented in Fig. 3, where proportions of the total variance of the pilot and navigator criteria are segregated and identified. The approach is

<sup>7</sup> In the AAF reports, this factor was designated as integration I, merely because it was found in some integration tests. The identity of integration I and memory IV still lacks direct verification. The tests saturated with integration I all require the remembering of rather complicated instructions.

exact, objective, and dependable. Others who use the same steps would be likely to arrive at similar results, though there may be differences of opinion as to naming and definition of some of the factors. The categories at least have referents. This type of analysis is possible for any job for which there are suitable criteria. The best procedure would be to correlate each criterion with a number of tests each of which features one common factor. Here, again, univocal tests would be most useful. Lacking the opportunity for this complete procedure, the job analyst could probably improve his inspectional methods materially by becoming well acquainted with the factors and using them as his reference categories. Once the factors become known, it is not very difficult to recognize by inspection their probable presence in either a new test or a new job criterion.

7. *Prediction of test validities and other intercorrelations.* When a job criterion becomes known in terms of common factors and their loadings, and when any new test has been analyzed and its loadings in corresponding factors have been determined, the validity of the test for predicting this criterion can be predicted by the use of equation (4). This statement should be modified by saying that the minimum validity can be estimated, allowing for the chance that there is unknown valid variance in the test. This feature of analytical methods became quite practical in the case of a few AAF tests that had not been correlated with the pilot criterion, and even more tests that had not been validated against the navigator criterion owing to the lack of opportunity. It is frequently feasible to gather the information needed for making a prediction of validity when it is not easy to go through the regular routine of validation by correlation of test with criterion. The experience of rea-

sonably close predictions with so many tests in the AAF lends some confidence to the practice of prediction, at least as a preliminary step. Later actual validation would be desirable.

What holds for the prediction of correlations between tests and criteria also holds for the intercorrelations among tests and among criteria. One would not need so often to forecast the correlation between two tests that had not been administered to the same population, though the possibility is real when each has been analyzed and a liberal portion of the non-error variance of each test has been identified. Of much more practical use is the prediction of the correlation between two criteria. An illustration of this was mentioned earlier, involving pilot and navigator criteria. Rarely does one have the opportunity to evaluate the same sample of individuals in two different jobs. Yet, for the sake of knowing how much differential prediction is possible, and how much independence to expect between the composite aptitude scores used to predict the two job criteria, some estimate of the intercorrelation is required. This estimate can be made using the common-factor loadings in equation (4).

8. *The assembly of test batteries.* From the knowledge of factorial composition of a criterion and of available tests, one can proceed to put together an appropriate battery with combining weights, with some assurance that one is working upon a rather dependable foundation. If one has used a collection of uni-factor tests in order to assess the factorial composition of the criterion, then of course a good battery could be selected from among their number. The advantages claimed here would be more apparent if it were desired to use tests that had not been previously correlated with the criterion but whose factorial compositions are



known. On a smaller scale, the advantage would be felt when it became necessary for any reason to substitute a new test for one in the battery. If the substituted test has a very similar factorial composition to that of the test it replaces, one can have considerable confidence that it will function in the same manner as the test replaced. This will be true even when the two tests superficially appear to be different. Without factorial knowledge, bad substitutions are undoubtedly made at times because it is assumed that superficial resemblances carry with them functional identities. Functional identities can be assured only through the duplication of factor variances.

#### SUMMARY

In this article it was pointed out that although tests that are developed on the basis of the work-sample principle are generally valid for predicting particular criteria which they resemble in a significant manner, they tend to be costly and of limited or unknown general applicability. Tests based upon the more common principles of job analysis are variously successful. Either when they are successful or when they are not, one is left very much in the dark as to the reasons, and progress is to a large degree of the trial-and-error type. Job-analysis categories are usually hazy in contour and seriously overlapping, thus introducing much wasteful effort. Tests developed to measure one supposed psychological entity are frequently found by factor analysis to be measuring something quite different. The factor-analysis approach to the problems of test development is proposed because it provides a rational, objective procedure and a meaningful, operationally defined, and dependable set of reference categories. Among its advantages the following are proposed for consideration:

1. It provides a precise segregation of what is measured in either a test or a job criterion into component variances.

2. It provides an economical procedure, eliminating wastage due to overlapping categories, and reducing the number of variables needed to encompass an enormous variety of individual differences.

3. It enables us to understand why a test is valid for prediction of behavior of a certain class and why it is not valid for prediction of behavior in some other class.

4. It leads to univocal or factorially pure tests, which are not only more economical for the coverage of aptitudes but are also more manageable in combinations and more meaningful in vocational guidance.

5. It leads to the discovery of new factors that are valid. The improvement of predictions depends upon this.

6. It provides an objective means of job analysis. Even if the opportunity does not exist for the statistical analysis of a job's requirements by correlation of its criteria with tests of known factorial composition, the knowledge and use of factorial categories will be a material aid to the observer of job activities.

7. Test validities may be predicted in advance of empirical validation procedures, where it is important to have this information early and when factorial compositions of both test and criterion are known. This kind of prediction also applies to the intercorrelations among tests and among criteria.

8. Knowledge of factorial composition of tests and of criteria is sufficient basis for the compilation of aptitude batteries. Knowing a job in terms of factors and their loadings, we can write a prescription of the tests and their weights needed to predict success in it.

In general terms, proceeding on the basis of factorial knowledge means

working with considerable light whereas test development or selection of tests without it means working in considerable darkness.

#### REFERENCES

1. *Army Air Forces Aviation Psychology Program Research Reports*. Reports No. 4-5. Washington, D. C.: Government Printing Office, 1947.
2. BENNETT, G. K., & FRY, D. E. *Mechanical comprehension test*. New York: Psychological Corporation.
3. GUILFORD, J. P. The discovery of aptitude and achievement variables. *Science*, 1947, 106, 279-282.
4. HULL, C. L. *Aptitude testing*. Yonkers: World Book Co., 1928.

# ON A DISTINCTION BETWEEN HYPOTHETICAL CONSTRUCTS AND INTERVENING VARIABLES

BY KENNETH MACCORQUODALE AND PAUL E. MEEHL

*University of Minnesota*

As the thinking of behavior theorists has become more sophisticated and self-conscious, there has been considerable discussion of the value and logical status of so-called 'intervening variables.' Hull speaks of "symbolic constructs, intervening variables, or hypothetical entities" (5, p. 22) and deals with them in his theoretical discussion as being roughly equivalent notions. At least, his exposition does not distinguish among them explicitly. In his presidential address on behavior at a choice point, Tolman inserts one of Hull's serial conditioning diagrams (11, p. 13) between the independent variables (maintenance schedule, goal object, etc.) and the dependent variable ('behavior ratio') to illustrate his concept of the intervening variable. This would seem to imply that Tolman views his 'intervening variables' as of the same character as Hull's. In view of this, it is somewhat surprising to discover that Skinner apparently feels that his formulations have a close affinity to those of Tolman, but are basically dissimilar to those of Hull (10, p. 436, 437). In advocating a theoretical structure which is 'descriptive' and 'positivistic,' he suggests that the model chosen by Hull (Newtonian mechanics) is not the most suitable model for purposes of behavior theory; and in general is critical of the whole postulate-deductive approach.

Simultaneously with these trends, one can still observe among 'tough-minded' psychologists the use of words such as 'unobservable' and 'hypothetical' in an essentially derogatory manner, and an almost compulsive fear of passing beyond the direct colligation of observable

data. 'Fictions' and 'hypothetical entities' are sometimes introduced into a discussion of theory with a degree of trepidation and apology quite unlike the freedom with which physicists talk about atoms, mesons, fields, and the like. There also seems to be a tendency to treat all hypothetical constructs as on the same footing merely because they are hypothetical; so that we find people arguing that if neutrons are admissible in physics, it must be admissible for us to talk about, *e.g.*, the damming up of libido and its reversion to earlier channels.

The view which theoretical psychologists take toward intervening variables and hypothetical constructs will of course profoundly influence the direction of theoretical thought. Furthermore, what *kinds* of hypothetical constructs we become accustomed to thinking about will have a considerable impact upon theory creation. The present paper aims to present what seems to us a major problem in the conceptualization of intervening variables, without claiming to offer a wholly satisfactory solution. Chiefly, it is our aim here to make a distinction between two subclasses of intervening variables, or we prefer to say, between 'intervening variables' and 'hypothetical constructs' which we feel is fundamental but is currently being neglected.

We shall begin with a common-sense distinction, and proceed later to formulations of this distinction which we hope will be more rigorous. Naively, it would seem that there is a difference in logical status between constructs which involve the hypothesization of an *entity*, *proc-*

*ess*, or *event* which is not itself observed, and constructs which do not involve such hypothesization. For example, Skinner's 'reflex reserve' is definable in terms of the total available responses without further conditioning, whereas Hull's 'afferent neural interaction' involves the notion of processes within the nervous system which presumably occur within the objective physical system and which, under suitable conditions, we might observe directly. To take examples from another science in which we psychologists may have less stake in the distinction, one might contrast the notion of 'resistance' in electricity to the notion of 'electron.' The resistance of a piece of wire is what Carnap has called a *dispositional concept*, and is defined by a special type of implication relation. When we say that the resistance of a wire is such-and-such, we mean that "so-and-so volts will give a current of so-and-so amperes." (For a more precise formulation of this see Carnap, 3, p. 440.) Resistance, in other words, is 'operational' in a very direct and primitive sense. The electron, on the other hand, is supposedly an *entity* of some sort. Statements about the electron are, to be sure, supported by means of observational sentences. Nevertheless, it is no longer maintained even by positivists that this set of supporting sentences exhaust the entire *meaning* of the sentences about the electron. Reichenbach, for example, distinguishes *abstracta* from *illata* (from Lat. *infero*). The latter are 'inferred things,' such as molecules, other people's minds, and so on. They are believed in on the basis of our impressions, but the sentences involving them, even those asserting their existence, are not reducible to sentences about impressions. This is the epistemological form, at rock bottom level, of the distinction we wish to make here.

The introduction of the word 'entity'

in our discussion has served merely to indicate the distinction, but in any crucial case there could be dispute as to whether a stated hypothesis involved the positing of an entity. For instance, is Hull's 'habit strength' an entity or not? Is 'drive' an entity? Is 'super-ego'?

Previous analyses of this difference may enable us to give a somewhat more precise formulation. These two kinds of concepts are variously distinguished by writers on philosophy of science. Feigl (personal communication) refers to *analytic* versus *existential* hypotheses. Benjamin (1) distinguishes between *abstractive* and *hypothetical* methods. In the abstractive or analytic method we merely neglect certain features of experience and group phenomena by a restricted set of properties into classes; relations between such classes can then be discovered empirically, and nothing has been added to the observed in the process. The hypothetical method, on the other hand, relates experiences "by inventing a fictitious substance or process or idea, in terms of which the experiences can be expressed. A hypothesis, in brief, correlates observations by adding something to them, while abstraction achieves the same end by subtracting something" (1, p. 184).

This quotation suggests to us at least three ways of stating the distinction we have in mind. First, it may be pointed out that in the statement of a hypothetical construction, as distinguished from an abstractive one, there occur words (other than the construct name itself) which are not explicitly defined by (or reduced to) the empirical relations. Once having set up sentences (postulates) containing these hypothetical words, we can arrive by deduction at empirical sentences which can themselves be tested. But the words themselves are not defined directly by or reducible to these empirical facts. This

is not true of abstractive concepts, such as resistance or solubility or, say, 'drive' as used by Skinner. (We may neglect wholly non-committal words such as *state*, which specify nothing except that the conditions are internal.)

A second apparent difference between abstractive and hypothetical concepts is in their logical relation to the facts, *i.e.*, the observation-sentences and empirical laws which are the basis for believing them. In the case of sentences containing only abstractive concepts, the truth of the empirical laws constitutes *both the necessary and sufficient conditions* for the truth of the abstractive sentences. For sentences involving hypothetical concepts, this is well known to be false. The empirical laws are necessary for the truth of the hypothetical sentences, since the latter imply them; but they are not sufficient. All scientific hypothesizing is in the invalid 'third figure' of the implicative syllogism. We neglect here the impossibility, emphasized by Reichenbach and others, of equating even an abstractive sentence or empirical 'law' to a *finite* number of particular observation sentences; this is of importance to philosophers of science but for help in the understanding of theories is of no particular consequence. We shall be assuming the trustworthiness of induction throughout and hence will treat 'direct' observational laws as universal sentences or as sentential functions. One can deduce empirical laws from sentences involving hypothetical constructs, but not conversely. Thus, beginning with the hypothesis that gases are made up of small particles which obey the laws of mechanics, plus certain approximating assumptions about the relation of their sizes to their distances, their perfect elasticity, and their lack of mutual attraction, one can apply mathematical rules and eventually, by direct substitution and equation,

lead without arbitrariness to the empirical equation  $PV = K$ . However, one cannot rigorously reverse the process. That is, one cannot commence with the empirical gas law  $PV = K$  and arrive at the full kinetic theory. The mathematics is reversible, granted that certain arbitrary breakups of constants *etc.*, are permitted; but beginning with the empirical law itself there is no basis for these arbitrary breakups. Furthermore, aside from the equations themselves, there are coordinated with these equations certain existence propositions, and assertions about the properties of the entities hypothesized. We state that there exist certain small particles, that they collide with the walls of the container, that the root mean square of their velocities is proportional to the temperature, *etc.* These assertions can of course not be deduced from the empirical law relating pressure and volume.

This suggests a third distinction between concepts of the two kinds. In the case of abstractive concepts, the quantitative form of the concept, *e.g.*, a measure of its 'amount,' can be derived directly from the empirical laws simply by grouping of terms. In the case of hypothetical concepts, mere grouping of terms is not sufficient. We are less assured of this distinction than of the other two, but we have not been able to think of any exceptions. It seems to us also that, in the case of Hull, this is the point which makes our distinction between hypothetical constructs and intervening variables most obvious. Let us therefore consider Hull's equations as an example.

In *Principles of Behavior*, the influence of certain independent variables such as number of reinforcements, delay in reward, stimulus-response asynchronism, *etc.*, upon response strength is experimentally investigated. In the study of the influence of each of these, the other independent variables are



held constant. The experimental findings lead to the formulation of the separate laws of dependence as a set of growth and decay functions. We shall neglect for the moment the complication of drive and of all other variables which intervene between the construct  ${}_sH_r$  and the empirical measure of response. That is to say, we shall deal only with the variables introduced in Hull's Postulate 4. The mathematical statement of Postulate 4 is

$${}_sH_r = M(1 - e^{-kw})e^{-ft}e^{-ut'}(1 - e^{-iN}).$$

(5, p. 178)

This equation asserts that habit strength is a certain joint function of four variables which refer to direct empirical quantities—number of reinforcements, delay in reinforcement, amount of reinforcement, and asynchronism between the discriminative stimuli and the response. It is important to see that in this case Hull does not distinguish the four experimentally separated laws combined in the equation by separate concept-names; the only intervening variable introduced is habit strength, which is written as an explicit function of four empirical variables  $w$ ,  $t$ ,  $t'$ , and  $N$ . It would be quite possible to introduce an intervening variable referring to, say, the last bracket only; it might be called 'cumulative reinforcement' and it would be a function of only one empirical variable,  $N$ . This would be the most reasonable breakdown of habit strength inasmuch as the other three growth functions (two negative) serve merely to modify the asymptote  $M$  (5, p. 181). That is to say, given a certain (maintained) rule for the amount of reinforcement given and two time-specifications concerning the constant relation of the response to two other operations, we have determined a parameter  $m$  for a dynamic curve describing the course of acquisition of habit strength. The quantity  $(1 - e^{-iN})$  (which

we are here calling 'cumulative reinforcement') is then an intervening variable which is multiplied by the parameter  $m$  in order to determine the value of habit strength after  $N$  reinforcements have occurred.

Suppose now that a critic asks us whether our 'cumulative reinforcement' really *exists*. This amounts to asking whether we have formulated a 'correct statement' concerning the relation of this intervening variable to the anchoring (empirical) variables. For since the statement of 'existence' for the intervening variable is so far confined to the equations above, the 'existence' of cumulative reinforcement reduces strictly to the second question. And this second question, as to whether the statement about the intervening variable's relation to the facts is correct, is in turn equivalent to the question, "Are the empirical variables related in such-and-such a way?" In other words, to confirm the equation for habit strength, it is merely necessary to state that (as Hull assumes in his earlier chapters) with drive, etc., constant, some empirical measure of response strength  $R$  is a linear function of habit strength. Then we can write directly,

$$R = C({}_sH_r) = C \cdot F(w)G(t)H(t')J(N) \\ = Q(w, t, t', N).$$

To confirm or disconfirm this equation is a direct empirical matter. It is possible to multiply out the bracketed quantities in various combinations, so as to make the arbitrary groupings disappear; what will mathematically persist through all such regroupings will be the rather complicated joint function  $Q$  of the four empirical variables  $w$ ,  $t$ ,  $t'$ , and  $N$ . By various arbitrary groupings and combinations we could define 15 alternative and equivalent sets of intervening variables. Thus, we might multiply out three of the four brackets in the basic equation but for some rea-

son choose to put  $e^{ut'}$  separately into the denominator. This would give us

$$R = \frac{F(w, t, N)}{e^{ut'}}$$

as the particular form for our empirical relation.  $F(w, t, N)$  could then be given an appropriate 'intervening variable' name, and the stimulus-response asynchronism  $t'$  would then define an intervening variable  $e^{ut'}$ .

It may be objected that 'habit strength' presumably refers to some state of the organism which is set up by reinforcing  $N$  times under specified conditions; whereas  $e^{ut'}$  cannot refer to any such state. This seems plausible; but the point is that to establish it as a state, it would be necessary to coordinate to the groupings within equations certain existence propositions, *i.e.*, propositions that do *more* than define a term by saying "Let the quantity  $G(x, y, z)$ , where  $x, y, z$  are empirical variables, be designated by the phrase so-and-so." This setting up of existence propositions could presumably be done even for a quantity such as  $e^{ut'}$ , by referring to such hypothetical processes as, say, diminishing traces in the neural reverberation circuits activated by a certain discriminative stimulus.

In the above example we have considered the fractionation of the intervening variable  $H_r$  into others. This reasoning can also be extended in the upward direction, *i.e.*, in the direction of fusion rather than fractionation. Let us treat 'habit strength' as Hull would treat our 'cumulative reinforcement,' by not giving it a name at all. It is still possible to set up equations to fit the Perin-Williams data (5, p. 229, 255) without referring to habit strength, writing merely

$$n = F(N, h),$$

where  $N$  and  $h$  are again both purely empirical variables.

We do not mean to imply that the divisions made by Hull (or Tolman) are of no value. It is convenient to have some term to refer to the result of a certain maintenance schedule, instead of having to say "that part of the general multivariable equation of response strength which contains '*hours since eating to satiety*' as an independent variable." We merely wish to emphasize that in the case of Hull's intervening variables, it is both necessary and sufficient for the truth of his 'theory' about the intervening variables that the empirical facts should be as his equations specify. The latter are merely names attached to certain convenient groupings of terms in his empirically fitted equations. It is always possible to coordinate to these quantities, which as written mathematically contain parameters and experimental variables only, certain existence propositions which would automatically make the construct 'hypothetical' rather than 'abstractive.' This giving of what Reichenbach calls 'surplus meaning' automatically destroys the equivalence between the empirical laws and the theoretical construct. When habit strength *means* the product of the four functions of  $w, t, t'$  and  $N$ , then if the response strength is related to these empirical variables in the way described, habit strength 'exists' in the trivial sense that the law holds. Our confidence in the 'correctness' of the intervening variable formulation is precisely as great as our confidence in the laws. When, however, habit strength means not merely this product of empirical functions but something more of a neural or other physiological nature, then the theory could be false even if the empirical relations hold.

It seems to us that Tolman himself, in using one of Hull's serial conditioning diagrams as a set of intervening variables (11, p. 13), departs from his

original definition. He has first described the situation in which the 'behavior ratio' is a complex function  $f_1$  of the independent experimental variables. He goes on to say,

"A theory, as I shall conceive it, is a set of intervening variables. These to-be-inserted intervening variables are 'constructs' which we, the theorists, evolve as a useful way of breaking down into more manageable form the original complete  $f_1$  function" (11, p. 9).

His reason for introducing intervening variables does not seem to us very cogent as he states it. He says that empirically establishing the form of  $f_1$  to cover the effects on behavior of all the permutations and combinations of the independent variables would be a 'humanly endless task.' If this means that all of the verifying instances of a continuous mathematical function cannot be empirically achieved it is true; but that is equally true for a function of one variable only. In order to utilize the proposed relationship between Tolman's function  $f_3$  (11, p. 10) which describes the relation of the behavior to the intervening variables, it is still necessary to establish empirically that the relationship holds—which amounts essentially to trying several of the infinitely many permutations and combinations (as in the Perin-Williams study) until we are inductively satisfied by ordinary scientific standards.

However cogent the arguments for intervening variables may be, it seems clear from Tolman's description that they are what we are calling *abstractive* rather than *hypothetical*. His notion of them involves nothing which is not in the empirical laws that support them. (We may speak of 'laws' here in the plural in spite of there being just the single function  $f_1$ , just as Boyle's and Charles' laws are distinguished in addition to the more general gas law  $PV/T = R$ .) For Tolman, the merit of

an intervening variable is of a purely 'summarizing' character. One can determine the function  $f_1$  by parts, so to speak (11, p. 17), so that the effect of a given maintenance schedule upon one part of  $f_1$  may be referred to conveniently as *drive*. For a given drive, we can expect such-and-such behavior ratios in a diversity of situations defined by various combinations of the other independent variables.

It has been observed earlier that in introducing one of Hull's well-known serial conditioning diagrams as an example of intervening variables outside Tolman's own system, we see a departure from the definition Tolman gives. The Hull diagrams contain symbols such as  $r_g$  (fractional anticipatory goal response) and  $s_g$  (the proprioceptive impulses produced by the movements constituting  $r_g$ ). These symbols refer to hypothetical processes within the organism, having an allegedly real although undetermined neuromuscular locus. These events are in principle directly observable. In fact, here the case for speaking of an objective reality is even stronger than Reichenbach's examples of electrons, molecules, etc.; since even the criterion of *technical* verifiability, admitted by all positivists to be too strong a restriction, would not exclude these hypotheses as empirically meaningless. Even without penetrating the organism's skin we have some direct observational evidence of  $r_g$  in the work of Miller (7). Whether  $r_g$  occurs and actually plays the role described is not relevant here; the point is that the diagrams and verbal explanations of Hull involve the supposition that it does. He assumes the existence of certain processes which are not logically implied by the empirical laws in the sense of strict equivalence. Even if, by using the notion of fractional anticipatory goal response, Hull deduced all of the

empirical laws relating independent and dependent variables, alternative hypotheses could be offered. Because of the 'surplus meaning' contained in concepts like  $r_g$  and  $s_g$ , these concepts are not really 'anchored' to the facts in the sense implied by Tolman's definition of intervening variables or by Hull's diagram on page 22 of the *Principles*. Hull states in reference to this diagram,

"When an intervening variable is thus securely anchored to observables on both sides it can be safely employed in scientific theory" (5, p. 22).

We presume that Hull means in this statement that the anchoring in question is not only a sufficient but a necessary condition for scientific admissibility. We feel that the criterion is too strong, assuming that the structure of modern physical science is to be allowed. This sort of anchoring makes the intervening variable strictly reducible to the empirical laws, which is, to be sure, what Tolman's original definition implied. But it excludes such extremely fruitful hypotheses as Hull's own fractional anticipatory goal responses, for which the strict reducibility does not exist.

It occurs to us also in this connection that Hull seems to have moved in the direction of Skinner and Tolman in his treatment of intervening variables. The use of the postulate-theorem approach is maintained more as a form in the *Principles* than as an actual instrument of discovery. In this respect, the *Principles* is much less like Hull's Newtonian model than was the *Mathematico-deductive theory of rote learning*. The justification of 'postulates' in the usual sense is their ability to mediate deductions of empirical laws which are then verified. In the *Principles*, the 'postulates' are verified directly, by the experimental device of holding all variables constant except the one for which

we want to find a law. This is quite unlike the derivation of the gas law in physics. The only sense in which any postulates are 'assumed' is in the assumption, referred to by Hull on page 181 of the *Principles*, that the separately verified parts of Postulate 4 will in fact operate according to his equation 16 when combined. This is certainly a 'postulate' only in a very attenuated sense, since it amounts essentially to an empirical extrapolation which can be verified directly, as Hull suggests.

At this point any distinction between the type of theory advocated by Hull and that advocated by Skinner or Tolman would seem to disappear, except for the relatively non-contributory 'neural' references contained in the verbal statement of Hull's postulates. Insofar as this neural reference is taken seriously, however, we are still dealing with concepts of a hypothetical rather than abstract character. There are various places in Hull's *Principles* where the verbal accompaniment of a concept, which in its mathematical form is an intervening variable in the strict (Tolman) sense, makes it a hypothetical construct. Thus, the operational definition of a *pav* of inhibition (5, p. 281) would seem merely to mean that when we know from the independent variables that the combined habit strength and drive, together with a discriminative stimulus located so many j.n.d.'s from the original, would yield a reaction potential of so many wats, it requires an equal number of *pavs* of inhibition to yield an effective reaction potential of zero. However, in the accompanying verbal discussion (5, p. 281) Hull refers to the removal of the inhibitory substance by the blood stream passing through effector organs as determining the quantitative law of spontaneous loss of inhibition as a function of time. 'Affluent neural interaction' is another ex-



ample of a concept which is mathematically represented as a relation of intervening variables in Tolman's sense, but to which are coordinated verbal statements that convey the surplus meaning and make it an hypothesis.

The question might be raised, whether this is not always the difference—that the mathematical assertions are definitive of intervening variables but the verbal additions lend the hypothetical character to such concepts. We do not believe this is the essential difference. There are mathematical expressions whose meaning is not defined in the absence of verbal existential accompaniment, because the quantities involved refer to non-observational (*i.e.*, hypothetical) processes or entities. There are other mathematical expressions for which this is not true, since their component symbols have direct observational reference. In the case of our 'cumulative reinforcement' term  $(1-e^{-iN})$ , no coordinated existential proposition is required. We simply say, "Response probability is such-and-such a multivariate function of such-and-such experimental variables. Within this function can be isolated a simple growth function of one variable, whose value as a function of  $N$  is referred to as *cumulative reinforcement*." This may be taken as an adequate reference for  $(1-e^{-iN})$ . On the other hand, in the derivation of the law  $PV = K$  there occur statements such as "When the gas is maintained at the same temperature,  $mv^2/2$  does not change." Neither  $m$  nor  $v$  is an empirical variable. This statement does not tell us anything *until* we are informed that  $v$  refers to the velocity which each molecule of the gas could be assumed to have in order that their mean kinetic energy should be what it is. In other words, in the derivation of the gas laws from kinetic theory there occur mathematical assertions whose meaning is unclear without

the accompanying existence assertions, and *which cannot be utilized to take the subsequent mathematical steps in the chain of inferences unless these assertions are included*. Thus, to get from a purely mathematical statement that a molecule on impact conserves all of its momentum, to a mathematical statement whose terms refer to the empirical concept of 'pressure on the walls,' it is necessary to know (from the accompanying verbal description) that in the equations of derivation,  $m$  refers to the mass of a hypothetical particle that strikes the wall,  $v$  to its velocity, and so on. This example shows that some mathematical formulations are themselves incomplete in the sense that they cannot mediate the desired deductions unless certain existential propositions are stated alongside, so as to render certain necessary substitutions and equations legitimate. Therefore it is not merely the matter of mathematical form that distinguishes a 'pure' intervening variable from a hypothesis.

In the second place, it seems to us that the use of verbal statements without mathematical formulations does not guarantee that we are dealing with a hypothetical construct rather than an intervening variable. Consider Skinner's definition of emotion as a 'state of the organism' which alters the proportionality between reserve and strength. This is not defined as a direct proportionality, and in fact Skinner nowhere deals with its quantitative form. No mathematical statement is given by him; yet we would contend that the use of the word 'state' does not in any way make the notion of emotion existential, any more than drive is existential in Skinner's usage. The 'state' of emotion is not to be described in any way except by specifying (a) The class of stimuli which are able to produce it and (b) The effects upon response strength. Hence emotion for Skinner is a true in-



tervening variable, in Tolman's original sense. We conclude from these examples that whether a given concept is abstractive or hypothetical is not merely a matter of whether it is an equation with or without accompanying verbal exposition.

On the basis of these considerations, we are inclined to propose a linguistic convention for psychological theorists which we feel will help to clarify discussion of these matters. We suggest that the phrase 'intervening variable' be restricted to the original use implied by Tolman's definition. Such a variable will then be simply a quantity obtained by a specified manipulation of the values of empirical variables; it will involve no hypothesis as to the existence of nonobserved entities or the occurrence of unobserved processes; it will contain, in its complete statement for all purposes of theory and prediction, no words which are not definable either explicitly or by reduction sentences in terms of the empirical variables; and the validity of empirical laws involving only observables will constitute both the necessary and sufficient conditions for the validity of the laws involving these intervening variables. Legitimate instances of such 'pure' intervening variables are Skinner's *reserve*, Tolman's *demand*, Hull's *habit strength*, and Lewin's *valence*. These constructs are the behavioral analogue of Carnap's 'dispositional concepts' such as solubility, resistance, inflammability, etc. It must be emphasized that the setting up of a definition or reduction for an intervening variable is not a wholly arbitrary and conventional matter. As Carnap has pointed out, it often happens that we give alternative sets of reduction sentences for the same dispositional concept; in these cases there is empirical content in our statement even though it has a form that suggests arbitrariness. The reason for this is that

these separate reductions for a given dispositional concept imply that the empirical events are themselves related in a certain way. The notion of amount of electric current can be introduced by several different observations, such as deposition of silver, deflection of a needle, hydrogen separated out of water, and so on. Such a set of reductions has empirical content because the empirical statements together with the reductions must not lead to contradictions. It is a contingent fact, not derivable from definitions alone, that the deposition of silver will give the same answer for 'amount of current' as will the deflection of a needle. A similar problem exists in Hull, when he sets up 'momentary effective reaction potential' as the last intervening variable in his chain. In the case of striated muscle reactions, it is stated that latency, resistance to extinction, and probability of occurrence of a response are all functions of reaction potential. Neglecting behavior oscillation, which does not occur in the formulation for the second two because they involve many repetitions of the situation, this means that the empirical variables must be perfectly correlated (non-linearly, of course). The only possible source of variation which could attenuate a perfect correlation between probability of occurrence and resistance to extinction would be actual errors of experimental measurement, since there are no sources of uncontrolled variation left within the organism. If we consider average latency instead of momentary latency (which is a function of *momentary* effective reaction potential and hence varies with behavioral oscillation), latency and resistance to extinction should also be perfectly correlated. It remains to be seen whether the fact will support Hull in giving simultaneously several reductions for the notion of reaction potential.

As a second linguistic convention, we

propose that the term 'hypothetical construct' be used to designate theoretical concepts which do *not* meet the requirements for intervening variables in the strict sense. That is to say, these constructs involve terms which are not wholly reducible to empirical terms; they refer to processes or entities that are not directly observed (although they need not be in principle unobservable); the mathematical expression of them cannot be formed simply by a suitable grouping of terms in a direct empirical equation; and the truth of the empirical laws involved is a necessary but not a sufficient condition for the truth of these conceptions. Examples of such constructs are Guthrie's M.P.S.'s, Hull's  $r_g$ 's,  $S_d$ 's, and *afferent neural interaction*, Allport's *biophysical traits*, Murray's *regnancies*, the notion of 'anxiety' as used by Mowrer, Miller, and Dollard and others of the Yale-derived group, and most theoretical constructs in psychoanalytic theory. Skinner and Tolman seem to be almost wholly free of hypothetical constructs, although when Skinner invokes such notions as the 'strain on the reserve' (10, p. 289) it is difficult to be sure.

We do not wish to seem to legislate usage, so that if the broader use of 'intervening variable' has become stuck in psychological discourse, we would propose alternatively a distinction between intervening variables of the 'abstractive' and of the 'hypothetical' kind. Since our personal preference is for restricting the phrase *intervening variables* to the pure type described by Tolman, we shall follow this convention in the remainder of the present paper.

The validity of intervening variables as we define them cannot be called into question except by an actual denial of the empirical facts. If, for example, Hull's proposed 'grand investigation' of the Perin-Williams type should be carried out and the complex hyperspatial

surface fitted adequately over a wide range of values (5, p. 181), it would be meaningless to reject the concept of 'habit strength' and still admit the empirical findings. For this reason, the only consideration which can be raised with respect to a given proposed intervening variable, when an initial defining or reduction equation is being written for it, is the question of convenience.

In the case of hypothetical constructs, this is not so clear. Science is pursued for many reasons, not the least of which is *n* *Cognizance*. Since hypothetical constructs assert the existence of entities and the occurrence of events not reducible to the observable, it would seem to some of us that it is the business of a hypothetical construct to be 'true.' It is possible to advance scientific knowledge by taking a completely 'as if' attitude toward such matters, but there are always those whose theoretical-cognitive need dictates that existential propositions should correspond to what is in fact the case. Contemporary philosophy of science, even as represented by those who have traditionally been most cautious about discussing 'truth' and most highly motivated to reduce it to the experiential, gives psychologists no right to be dogmatic about the 'as if' interpretation of theoretical knowledge (*cf.* especially Carnap, 4, p. 598, Kaufmann, 6, p. 35, Russell, 9, Introduction and Chapter XXI, and Reichenbach, 8, *passim*). We would find it rather difficult to defend the ingenious conditioning hypotheses developed in Hull's series of brilliant papers (1929-) in the *PSYCHOLOGICAL REVIEW* on the ground that they merely provide a "convenient shorthand summarization of the facts" or are of value in the 'practical manipulation' of the rat's behavior. We suspect that Professor Hull himself was motivated to write these articles because he considered that the hypothetical events represented in his diagrams

may have actually *occurred* and that the occurrence of these events represents the underlying truth about the learning phenomena he dealt with. In terms of practical application, much (if not most) of theoretical psychology is of little value. If we exclude the interesting anecdotes of Guthrie, contemporary learning theory is not of much use to school teachers. As a *theoretical* enterprise, it may fairly be demanded of a theory of learning that those elements which are 'hypothetical' in the present sense have some probability of being in correspondence with the actual events underlying the behavior phenomena, *i.e.*, that the assertions about hypothetical constructs be true.<sup>1</sup>

Another consideration may be introduced here from the standpoint of future developments in scientific integration. Even those of us who advocate the pursuit of behavioral knowledge on its own level and for its own sake must recognize that some day the 'pyramid of the sciences' will presumably catch up with us. For Skinner, this is of no consequence, since his consistent use of intervening variables in the strict sense genuinely frees him from neurophysiology and in fact makes it possible for him to impose certain conditions upon

neurophysiological explanations (10, pp. 429-431). Since he hypothesizes nothing about the character of the inner events, no finding about the inner events could prove disturbing to him. At most, he would be able to say that a given discovery of internal processes must not be complete because it cannot come to terms with his (empirical) laws. But for those theorists who do not confine themselves to intervening variables in the strict sense, neurology will some day become relevant. For this reason it is perhaps legitimate, even now, to require of a hypothetical construct that it should not be manifestly unreal in the sense that it assumes inner events that cannot conceivably occur. The 'as if' kinds of argument sometimes heard from more sophisticated exponents of psychoanalytic views often seem to ignore this consideration. A concept like *libido* or *ego* or *super-ego* may be introduced initially as though it is to be an intervening variable; or even less, it is treated as a merely conventional designation for a class of observable properties or occurrences. But somewhere in the course of theoretical discussion, we find that these words are being used as hypothetical constructs instead. We find that the *libido* has acquired certain hydraulic properties, or as in Freud's former view, that the 'energy' of *libido* has been converted into 'anxiety.' What began as a name for an intervening variable is finally a name for a 'something' which has a host of causal properties. These properties are not made explicit initially, but it is clear that the concept is to be used in an explanatory way which requires that the properties exist. Thus, *libido* may be introduced by an innocuous definition in terms of the 'set of sexual needs' or a 'general term for basic strivings.' But subsequently we find that certain puzzling phenomena are *deduced* ('explained') by means of the various prop-

<sup>1</sup> It is perhaps unnecessary to add that in adopting this position we do not mean to defend any form of metaphysical realist thesis. The ultimate 'reality' of the world in general is not the issue here; the point is merely that the reality of hypothetical constructs like the atom, from the standpoint of their logical relation to grounds, is not essentially different from that attributed to stones, chairs, other people, and the like. When we say that hypothetical constructs involve the notion of 'objective existence' of actual processes and entities within the organism, we mean the same sort of objective existence, defined by the same ordinary criteria, that is meant when we talk about the objective existence of Singapore. The present discussion operates within the common framework of empirical science and common sense and is intended to be metaphysically neutral.

erties of libido, e.g., that it flows, is dammed up, is converted into something else, tends to regress to earlier channels, adheres to things, makes its 'energy' available to the ego, and so on. It is naive to object to such formulations simply on the ground that they refer to unobservables, or are 'hypothetical,' or are not 'statistical.' None of these objections is a crucial one for any scientific construct, and if such criteria were applied a large and useful amount of modern science would have to be abandoned. The fundamental difficulty with such theories is two-fold. First, as has been implied by our remarks, there is the failure explicitly to announce the postulates concerning existential properties, so that these are introduced more or less surreptitiously and *ad hoc* as occasion demands. Secondly, by this device there is subtly achieved a transition from admissible intervening variables to inadmissible hypothetical constructs. These hypothetical constructs, unlike intervening variables, are inadmissible because they require the existence of entities and the occurrence of processes which cannot be seriously believed because of other knowledge.

In the case of libido, for instance, we may use such a term legitimately as a generic name for a class of empirical events or properties, or as an intervening variable. But the allied sciences of anatomy and physiology impose restrictions upon our use of it as a hypothetical construct. Even admitting the immature state of neurophysiology in terms of its relation to complex behavior, it must be clear that the central nervous-system does not in fact contain pipes or tubes with fluid in them, and there are no known properties of nervous tissue to which the hydraulic properties of libido could correspond. Hence, this part of a theory about 'inner events' is likely to remain metaphorical. For a

genuine intervening variable, there is no metaphor because all is merely shorthand summarization. For hypothetical constructs, there is a surplus meaning that is existential. We would argue that dynamic explanations utilizing hypothetical constructs ought not to be of such a character that they *have* to remain only metaphors.

Of course, this judgment in itself involves a 'best guess' about the future. A hypothetical construct which seems inherently metaphorical may involve a set of properties to which hitherto undiscovered characteristics of the nervous system correspond. So long as the propositions about the construct are not stated in the *terms* of the next lower discipline, it is always a possibility that the purely formal or relational content of the construct will find an isomorphism in such characteristics. For scientific theories this is enough, since here, as in physics, the associated mechanical imagery of the theorist is irrelevant. The tentative rejection of libido would then be based upon the belief that no neural process is likely to have the *combination* of formal properties required. Strictly speaking, this is always problematic when the basic science is incomplete.<sup>2</sup>

#### SUMMARY

1. At present the phrases 'intervening variable' and 'hypothetical construct' are often used interchangeably, and theoretical discourse often fails to distinguish what we believe are two rather different notions. We suggest that a failure to separate these leads to fundamental confusions. The distinction is between constructs which merely abstract the empirical relationships (Tolman's original intervening variables) and those constructs which are 'hypothetical' (*i.e.*, involve the supposition of

<sup>2</sup> We are indebted to Dr. Herbert Feigl for a clarification of this point.



entities or processes not among the observed).

2. Concepts of the first sort seem to be identifiable by three characteristics. First, the statement of such a concept does not contain any words which are not reducible to the empirical laws. Second, the validity of the empirical laws is both necessary and sufficient for the 'correctness' of the statements about the concept. Third, the quantitative expression of the concept can be obtained without mediate inference by suitable groupings of terms in the quantitative empirical laws.

3. Concepts of the second sort do not fulfil any of these three conditions. Their formulation involves words not wholly reducible to the words in the empirical laws; the validity of the empirical laws is not a sufficient condition for the truth of the concept, inasmuch as it contains surplus meaning; and the quantitative form of the concept is not obtainable simply by grouping empirical terms and functions.

4. We propose a linguistic convention in the interest of clarity: that the phrase *intervening variable* be restricted to concepts of the first kind, in harmony with Tolman's original definition; and that the phrase *hypothetical construct* be used for those of the second kind.

5. It is suggested that the only rule for proper intervening variables is that of convenience, since they have no factual content surplus to the empirical functions they serve to summarize.

6. In the case of hypothetical constructs, they have a cognitive, factual reference in addition to the empirical data which constitute their support. Hence, they ought to be held to a more stringent requirement in so far as our interests are theoretical. Their actual existence should be compatible with general knowledge and particularly with whatever relevant knowledge exists at the next lower level in the explanatory hierarchy.

#### REFERENCES

1. BENJAMIN, A. C. *An introduction to the philosophy of science*. New York: Macmillan, 1937.
2. CARNAP, R. Testability and meaning, Parts I-III. *Phil. Sci.*, 1936, 3, 419-471.
3. CARNAP, R. Testability and meaning, Part IV. *Phil. Sci.*, 1937, 4, 1-40.
4. —. Remarks on induction and truth. *Phil. & phenomenol. res.*, 1946, 6, 590-602.
5. HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
6. KAUFMANN, F. *Methodology in the social sciences*. London: Oxford University Press, 1944.
7. MILLER, N. E. A reply to 'Sign-Gestalt or conditioned reflex.' *PSYCHOL. REV.*, 1935, 42, 280-292.
8. REICHENBACH, H. *Experience and prediction*. Chicago: University of Chicago Press, 1938.
9. RUSSELL, B. *Inquiry into meaning and truth*. New York: Norton, 1940.
10. SKINNER, B. F. *Behavior of organisms*. New York: Appleton-Century, 1938.
11. TOLMAN, E. C. The determiners of behavior at a choice point. *PSYCHOL. REV.*, 1938, 45, 1-41.



# REACTION TO FRUSTRATION—A CRITIQUE AND HYPOTHESIS<sup>1</sup>

BY S. STANSFELD SARGENT

*Barnard College, Columbia University*

The problem of frustration has commanded considerable attention during the last decade. Not only psychiatrists and specialists in clinical and abnormal psychology, but also students of personality and social psychology have become interested in frustration. The *Frustration and Aggression* volume by the Yale collaborators, Dollard, Doob, Miller, Mowrer and Sears (1), has helped focus attention on the subject, as have papers by Maslow (4, 5), Rosenzweig (7, 8) and others.

While dealing with the concept of frustration in a course in Social Psychology, I became convinced that current treatments of frustration still lack a systematic framework and a clear definition of terms. The major concepts which are used lack integration; for example, frustration, conflict, motives, emotions, defense mechanisms, habit patterns, personality factors of many kinds, and situational influences. Several instances might be cited to illustrate confusion in usage of terms. 'Frustration' usually refers to environmental blocking of motives, but sometimes to an unpleasant emotional state resulting from the blocking. At times 'hostility' seems to mean actual behavior; again it signifies a strong feeling underlying behavior. 'Inferiority,' 'insecurity,' 'anxiety,' 'guilt' and many other concepts are frequently employed in ways which are unclear psychologically. Probably the worst of all is 'aggression,' which

sometimes seems to mean a motive, sometimes an emotional state akin to anger, sometimes a habit of mechanism, and sometimes a type of overt behavior!

I wish to propose a rather simple conceptual scheme for describing behavior resulting from frustration. It is presented as a hypothesis which seems reasonably consistent with clinical and experimental data and also with many of the theoretical formulations which have been advanced.

Briefly the hypothesis is this: frustration evokes a patterned sequence of behavior whose chief stages or aspects are indicated by the terms *frustration*, *emotion*, *habit or mechanism*, and *overt behavior*. The nature of each stage of the total process is determined by the interaction of two major factors: the individual's past experience, and the present situation as perceived or defined by the individual. Let us consider each of these in more detail.

It is well agreed that frustration involves the thwarting or blocking of a person's dominant motives, needs, drives, desires or purposes. However, some psychologists place greater stress upon the thwarting than upon the individual's reaction to it. For example, the Yale group defines frustration as "that condition which exists when a goal-response suffers interference" (1, p. 11). In his recent book Symonds defines it as "the blocking or interference of the satisfaction of an aroused need through some barrier or obstruction" (12, p. 51). Others emphasize not so much the thwarting, *per se*, as the significance of the thwarting to the individual. Maslow (5) insists that frustration involves

<sup>1</sup> This article is a slightly revised form of a paper presented at the 1946 meetings of the Eastern Psychological Association. I wish to thank Professor R. S. Woodworth for several helpful criticisms and suggestions.

two concepts—deprivation, and threat to the personality. Sexual deprivation, for example, does not necessarily constitute frustration, but when such deprivation is felt by the individual to represent rejection by the opposite sex, inferiority, or lack of respect, it becomes seriously frustrating. Similarly, Rosenzweig (7) distinguishes between 'need-persistent' and 'ego-defensive' reactions, the latter representing greater frustration. Zander (15) maintains that frustration occurs only when there is interference with "a goal believed important and attainable by a given person." In all probability future studies of frustration will take into account such subjective individual differences as are mentioned by Maslow, Rosenzweig and Zander.

In any event, we turn next to the question, What is the immediate psychological consequence of frustration? It is definitely not aggression, as most readers of *Frustration and Aggression* might assume. Nor is it the adoption of some handy defense mechanism, as others might conclude. First in time, and foremost in significance, frustration arouses a *pronounced emotional reaction*.

Most students of frustration refer to concomitant emotional tensions, but they seldom make emotion a central aspect of the whole reaction pattern.<sup>2</sup> According to the present hypothesis, emotion is the core of reaction to frustration. If no emotion is aroused, there is no frustration—at least not in any psychologically meaningful sense.

Furthermore, the emotion aroused may be broad and diffuse, like a generalized anger or fear, or it may be fairly specific, like hostility, jealousy, inferiority or shame. Whether the emotion is general or specific depends largely upon the nature of the whole precipitating

situation as interpreted by the individual.

It is clearly established that strong emotional reactions upset the organism and tend to pass over into overt behavior. However, the form of the resultant behavior is not, *ipso facto*, determined by the kind and intensity of the emotion. Behavior is, of course, partly dependent upon the emotion which agitates the organism; anger is more likely to work itself out in aggressive behavior than is anxiety or shame. But the form of the overt reaction is importantly affected by the individual's adjustive habits or mechanisms, and by the way he interprets the situation.

The above analysis agrees rather well with Rosenzweig's interpretation (8). In studying reactions to frustration, according to Rosenzweig, we must be concerned not with what is objectively present, but instead with what the individual emphasizes or reads into the situation according to his personality needs and traits. He finds three main types of reaction to frustration. The 'extrapunitive' is an aggressive reaction toward others. It arises from anger and indignation and from the individual's judgment which blames others; "I'll get you!" is its thesis. Thus, if snubbed by a friend, the extrapunitive reaction is to regard him as ill-bred and ungrateful. The 'intropunitive' is an aggressive reaction directed toward the self. It comes from feelings of humiliation and guilt, and from judgments of self-blame. The intropunitive reaction to a snub is to regard oneself as inferior and unworthy. The 'impunitive' reaction is unaggressive. It arises from feelings of embarrassment and shame and from the judgment "It can't be helped." A friend's snub would be condoned or glossed over as an oversight.

More than any other interpreter of frustration, Rosenzweig stresses the importance of both emotional and 'apper-

<sup>2</sup> Maslow and Rosenzweig are exceptions to this statement.

ceptive' or judgmental factors. I feel, however, that he has made the latter too conscious. According to my hypothesis there is a continuously operating, relatively unconscious perceptual process which may be called 'defining the situation.'

This term is taken from the sociologist, W. I. Thomas (13). It was used by him and by others (*e.g.*, Waller [14]) to designate the process of perceiving and interpreting, and also of exploring the behavior possibilities of a social situation. It has elements in common with Lewin's 'psychological environment' (3) and with Sherif's 'frames of reference' (11).<sup>3</sup> But 'defining the situation' is more than perceiving; it is a kind of active perceiving, interpreting and sizing up a situation with reference to one's potential behavior in it. We cannot know how a given situation influences an individual unless we know how he defines it for himself.

Strong emotions, then, tend toward overt behavior, but always directed and limited by the individual's adjustive habits and by the way he defines the situation. He may customarily express his emotions freely, or he may repress them. Or he may be adept at utilizing substitute forms—*i.e.*, mechanisms—for expressing his strong emotions which are the essence of frustration. Generally speaking, the more stress or threat he reads into the immediate social situation, the more inhibited and disguised his expressive behavior will be.

Our analysis will be made clearer by the use of an example and a diagram. (See Fig. 1.) An individual intent upon an important promotion in his business or profession learns the promotion has been won by another per-

son, which produces real frustration. He becomes emotional—but how? If the event is unexpected and the cause unclear, the emotion is a generalized sort of anger. If he knows, or thinks he knows, whose efforts defeated him, his emotional reaction takes the more specific form of hostility or hatred, quite possibly with components of jealousy. Psychologically this is a different phenomenon from generalized anger (though it may be similar physiologically) since it is directed toward a particular individual.

Let us assume, however, that our individual has no detailed information about the events leading up to the loss of his expected promotion and that, therefore, he is in a state of generalized anger. Then what? If he characteristically expresses emotion in an uninhibited way, he may throw things, kick chairs around and curse vehemently. But he is less likely to do this if persons whose opinions he values are present. If they are, he might rather engage in some substitute type of expression, such as rationalizing or seeking sympathy.

On the other hand, he may be the kind of person who seldom gives free vent to his emotions. He may then displace his anger upon his wife and children if they are present. He might kick the dog or cat, or 'take it out' on a clumsy delivery boy, all depending upon who is present at the time and what his relationship to them happens to be. Or if he were a person of violent prejudices, he might displace his anger upon 'the Jews,' 'the Reds,' 'the Catholics' or some other handy scapegoat. Again, he might regress; if his mother were present he might burst into tears and put his face in her lap as he always did when a child. Or he might engage in one or another kind of comforting fantasy.

Actually he would probably utilize more than one kind of defense mecha-

<sup>3</sup> See Sargent, S. S., *Defining the situation—an essential concept for social psychology*. Forthcoming publication based upon a paper delivered at the 1947 meetings of the Eastern Psychological Association.

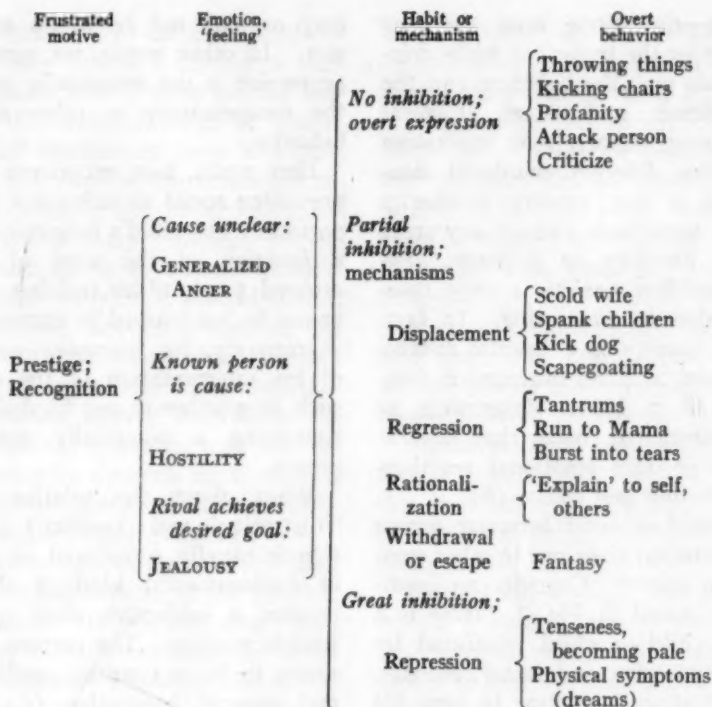


FIG. 1. Reaction to frustration.

nism. Seldom does a single outlet relieve all of one's strong emotional tensions. An immediate emotional outburst might well be followed by rationalizing, fantasy, or some kind of compensatory behavior. Clinical data suggest that as children most of us acquire quite a repertory of forms of substitute expression. Hence the particular one or ones we employ depend in large measure upon the social situation as we interpret it.

Another possibility is that, because of past training and/or a very stringent social situation, an individual may inhibit or repress nearly all overt behavior. If so, we would expect some sort of delayed overt expression, possibly in disguised form, as in dreams or physical symptoms of illness.

Does frustration eventuate in aggression? At the beginning of their book

the Yale psychologists propose "that the existence of frustration always leads to some form of aggression" (1, p. 1). This thesis is hard to defend, as two of the authors, Miller (6) and Sears (9), point out in subsequent articles.<sup>4</sup>

Much behavior resulting from frustration is, of course, aggressive. Probably the Yale group arrived at their sweeping conclusion partly because the cases they considered were dramatic, short-time, anger-producing kinds of

<sup>4</sup> The first part of the same proposition is "that the occurrence of aggressive behavior always presupposes the existence of frustration." We shall not discuss the subject here, except to suggest that it is also difficult to defend as a general statement. Certain kinds of behavior which are definitely aggressive seem to be the socially sanctioned ways of behaving in some communities (e.g., a tough city slum area or a primitive culture). Such behavior may well be learned and practiced without having its origin, necessarily, in frustration.



frustration—the young man who was bawled out by the traffic cop while driving with his girl, the boarders and the delayed dinner, and so on. Clinical data, however, suggest that frustration may produce different emotional reactions, such as fear, anxiety, inferiority or shame, sometimes without any trace of anger, hostility or jealousy. Symonds considers anxiety a very common reaction to frustration. In fact, he defines anxiety as a “mental distress with respect to some anticipated frustration” (12, p. 133). Rosenzweig, as already mentioned, notes that frustration may produce emotional reactions like humiliation and shame (8).

What kind of overt behavior occurs when frustration gives rise to other emotions than anger? Consider an example, diagrammed in Fig. 2. Here is a ‘rejected’ child—a child frustrated by denial of affection and social response. If the situation is unclear to him, his emotional reaction is one of general anxiety or insecurity. What does he do? He may, indeed, compensate and engage in bullying, boasting and other kinds of aggressive behavior. But he may instead compensate by making a friend of the teacher or by forming a strong attachment for an older boy or girl. He may seek satisfaction through identification—either by playing the role of a ‘big shot,’ or by joining some social group with prestige value. If he withdraws and daydreams, his fantasies

may or may not be of an aggressive sort. In other words, we suggest that aggression is not necessarily present in the compensatory or other substitute behavior.

Here again, past experience and the prevailing social situation are both important. The child’s behavior is partly a function of the kind of emotion aroused, partly of his training—whether or not he has learned to express himself aggressively, for example,—and partly of his interpretation of the situation, such as whether or not he defines it as containing a potentially sympathetic person.

What about the relation between ‘frustration’ and ‘conflict’? Frustration is usually considered an objective or environmental kind of thwarting; conflict a subjective clash of incompatible motives. The current tendency seems to be to consider conflict a special case of frustration (e.g., Shaffer [10, p. 118] and Rosenzweig [7]). Many psychologists, however, treat frustration and conflict separately and do not attempt to relate them. In terms of the present hypothesis the important point is that both frustration and conflict involve dynamic and highly upsetting emotional states which impel the organism toward some sort of overt behavior. Reaction to conflict, as to frustration, follows the same sequence: emotion, habit or mechanism, and overt behavior. For instance, conflict arising

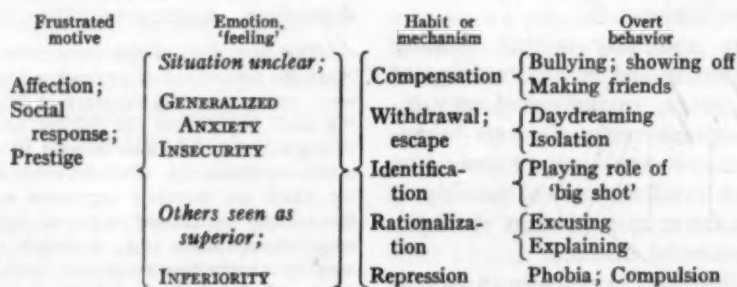


FIG. 2. Reaction to frustration.



from performance of an act considered immoral may arouse a feeling of generalized anxiety, or a more specific feeling of guilt, embarrassment or shame. Habits and mechanisms come into play. Through sublimation the emotional reaction may impel one toward religious or altruistic activity; through projection, toward gossip or scandal-mongering; through repression, toward phobia or compulsion; through a kind of displacement toward masochism or other self-directed aggression. The nature of the frustration largely determines the basic emotional reaction, and the resultant behavior depends upon existing habit-patterns operating in the individually defined social situation.

The above discussion has omitted many important aspects of the problem of frustration. It has not dealt with the efficacy of resultant behavior in reducing emotional tensions evoked by frustration. It has neglected the important matters, so ably treated by the Yale group, of the effects of differing degrees and strengths of instigation, or the effects of anticipated punishment. Nor has it dealt with the concept of 'frustration tolerance' which is taken up by Rosenzweig (7) and others.

The critique and hypothesis presented above is a systematic contribution designed to fill in certain gaps and to fit loose ends together. Some psychologists may object to such an analysis, *per se*, as violating the essential unity or Gestalt-character of behavior. The only answer, I suppose, is that some kinds of behavior are too complex to be treated as a whole; they have to be analyzed, though efforts must be made to put the pieces together again. Other psychologists will undoubtedly object to certain of the statements and interpretations. The whole hypothesis needs, of course, to be verified by clinical or experimental methods.

In addition to setting up the four-

stage scheme of frustration-emotion-mechanism-overt behavior, the hypothesis proposes the following things:

It makes emotion the central dynamic factor in reaction to frustration, and distinguishes between generalized emotional states (e.g., anger, anxiety) and more specific and directed states or 'feelings' (e.g., hostility, jealousy, inferiority).

It stresses the interoperation of both past experience and present situations as determining the form and content of resulting overt behavior.

Furthermore, it emphasizes that the crucial present factor is not the situation as it exists in some objective sense, but rather as the individual defines and interprets it.

Most of all, perhaps, this paper represents a protest against what Leeper calls 'peripheralism' in psychology (2); that is to say, the description of behavior chiefly in terms of stimuli and overt responses, to the neglect of intervening organismic factors. Hence it is, in brief, an attempt to describe all the significant psychological variables which interoperate when a person is frustrated.

#### BIBLIOGRAPHY

1. DOLLARD, J., DOOB, L. W., MILLER, N. E., MOWRER, O. H., & SEARS, R. R. *Frustration and aggression*. New Haven: Yale Univ. Press, 1939.
2. LEEPER, R. The experimental psychologists as reluctant dragons. *Amer. Psychol.*, 1946, 1, 295.
3. LEWIN, K. *Dynamic theory of personality*. New York & London: McGraw-Hill Book Co., 1935.
4. MASLOW, A. H. Conflict, frustration and the theory of threat. *J. abnorm. & soc. Psychol.*, 1943, 38, 81-86.
5. ——. Deprivation, threat and frustration. *PSYCHOL. REV.*, 1941, 48, 364-366.
6. MILLER, N. E. The frustration-aggression hypothesis. *PSYCHOL. REV.*, 1941, 48, 337-342.
7. ROSENZWEIG, S. An outline of frustration theory. In *Personality and the behav-*

- ior disorders (J. McV. Hunt, Ed.). New York: Ronald Press, 1944, Chap. 11.
8. —. Types of reaction to frustration. *J. abnorm. & soc. Psychol.*, 1934, 29, 298-300.
  9. SEARS, R. R. Non-aggressive reactions to frustration. *PSYCHOL. REV.*, 1941, 48, 343-346.
  10. SHAFFER, L. F. *The psychology of adjustment*. Boston & New York: Houghton Mifflin, 1936.
  11. SHERIF, M. *The psychology of social norms*. New York & London: Harper & Bros., 1936.
  12. SYMONDS, P. M. *The dynamics of human adjustment*. New York: American Book Co., 1946.
  13. THOMAS, W. I. *The unadjusted girl*. Boston: Little, Brown, 1923.
  14. WALLER, W. W. *The sociology of teaching*. New York: John Wiley, 1932.
  15. ZANDER, A. F. A study of experimental frustration. *Psychol. Monogr.*, 1944, 56, No. 256.

# RELATIONS BETWEEN PHILOSOPHY AND PSYCHOLOGY<sup>1</sup>

BY ALBERT G. A. BALZ

*University of Virginia*

What relations obtain between philosophy and psychology? More importantly, in what ways, if any, are they uniquely related? The general question could be asked: What relations obtain between the sciences, severally and as a whole, and philosophical thought? Is the question concerning psychology and philosophy merely a special case of the more inclusive one? Or has psychology an especial relevance to philosophical thought? Is it peculiarly contributive to philosophy? Is philosophy peculiarly dependent upon psychology? The philosophic tradition, at least in modern times, suggests that it is. If this be the case, just why is it so? These questions could be reversed. It is, of course, the hope of the philosopher that his reflection contributes to the pursuit of scientific inquiry in every field. Does philosophy, however, contribute to psychology in an exceptional way and to a special degree? In sum, is psychology dependent upon philosophical thought to an extent not holding with respect to other sciences? Is there a peculiarly important inter-dependence?

Many psychologists, one suspects, would deny the dependence of the science upon philosophy, even though admitting philosophy's dependence upon psychology. They might declare that the influence of philosophy, or at any rate of some philosophies and some philosophical ideas, upon psychology has been deplorable. It has retarded the development of the science. Philo-

sophical ideas have exerted an alien control upon psychological inquiry. It is better, some psychologists may insist, to make the science as independent of philosophy as is now the case with astronomy, chemistry, or entomology. If one could be sure that such contentions are sound, discussion could be limited to the question of the dependency of philosophy upon psychology—and the question might have little interest for the psychologist.

It could be urged, in general, that the questions raised cannot be dealt with fruitfully, at least at the present time. There is no assured body of philosophic doctrine, commanding general acceptance, it may be contended, as there are such bodies of scientific results. There are only many philosophies, more or less systematic, and many philosophers, with as many more or less confused standpoints. The dependency of philosophy upon psychology, accordingly, may be conceived in as many ways as there are such standpoints. In these conditions, moreover, little meaning can be given to the notion of psychology's dependence upon philosophy. On the other hand, psychology itself has many subdivisions. There is no consensus as to the bases of psychological inquiries or their objectives. If one could screen away from physics, say, everything that is at present controversial, conflicting and unsettled, there would remain an impressive residue of verified findings. With respect to philosophy, a comparable screening process would leave no residue—save perhaps the fact that everything in philosophy is controversial. What would remain if psychology were screened? Would we discover an im-

<sup>1</sup> This paper was read, in abbreviated form, before the Southern Society for Philosophy and Psychology, thirty-ninth annual meeting, Washington University, St. Louis, April 3-5, 1947.

pressive organized residue, an assured system of truths, so that we might refer to the science of psychology as we refer to the science of physics? If this be the case, and if philosophic thought be genuinely dependent upon psychology, the inchoate state of philosophy appears strange. Why have not philosophers profited by the findings of psychology with the result that beneath controversy can be discerned a body of universally accepted doctrines? Or can it be that psychology possesses little unity, a scanty fund of established results, and therefore has little to contribute to philosophy? Perhaps indignant psychologists, grudgingly admitting that psychology is scarcely more than a name for scores of semi-independent enterprises, may retort that this condition is due to the unhappy influence of philosophies, if not of philosophy. All in all, it could be contended with show of reason that there is little to be gained by discussing the relations of philosophy and psychology.

I propose to offer a series of conjectures in the hope that some approximation to mutual understanding may ensue. I propose further to take advantage of Plato's rhetoric. It may be that I shall also be adopting or adapting something of Plato's thought, but this may be left to one side.

Plato said something to the effect that the philosopher is the spectator of all time and existence. To endeavor to be the spectator of all time and existence, then, indicates one function of the philosopher. He must try to assume a universal spectatorial point of view. This could be called his cosmological function. Let us accept this as a first approximation concerning philosophy. A corresponding first approximation with respect to psychology is in order. What does psychology undertake to do? It appears to be peculiarly concerned with and about man. There are many sci-

ences concerned with man. Such sciences as biology and chemistry comprise man within the scope of what is called their subject-matter. But man is within their subject-matter incidentally. As types of inquiry, supplied with a fund of basic ideas, they can be focussed upon man, with the result that sub-sciences arise. The matter need not be pursued into detail. It seems clear that no other science is concerned with man so concretely, so centrally, so eminently, as psychology. Such sciences as anthropology or sociology would appear to be, in theory at least, off-shoots or applications of psychology. Presumably their success would depend upon that of psychology. It is the case that psychology does, and perhaps must, take advantage of other sciences. But it does this for the sake of realizing its own objective, and this, as a first approximation, is knowledge of the nature of man and of all that depends upon that nature. Indeed, the interpretation of the real significance of the findings of many other sciences is conditioned by psychology. This science, then, may be described as the preeminent anthropocentric science. It is first of all about man. It may also be about the universe viewed in its relation to man. For psychology, accordingly, man is central in all time and existence.

The question may now be posed: In relation to the spectatorial function of philosophy, what is the peculiar significance and contribution of psychology? It may be conceded that the cosmological function, if and when exercised, is exercised by men. But the function as such requires that this fact be neglected. As a matter of procedure, the spectatorial function must be carried on *as if* what *is* the case, that it is a functioning of man, were *not* the case. At any rate, and in some respects, the fact must be regarded as irrelevant to the effort to view all time and existence. Presum-



ably, the exercise of the function has an objective. What could this be save knowledge of all time and existence, knowledge of time and of what is unrolled in time? We must be restrained in our conjectures. The spectatorial function, in its rhetoric at least, suggests that there *may* be a knowledge of all time and existence, and that all time and existence *may* be more or less of a whole, more or less of a single ordered realm, marked by widely pervasive relations. It is not claimed that there *is* a whole of being, but only that there *may* be. It is not urged that the spectatorial function does view the whole of being, if there is one, but only that this effort is defined by the possibility. If it is assumed that this is one function of philosophy, if it is assumed that the effort is to view everything that is displayed in time, what is to be viewed may be signified by a single word, with a capital letter, *i.e.*, Nature. If nature is in some degree and measure *one*, in some degree and measure an ordered domain, then in so far as it is permeated by order it may be said to have a Logos. So far as the Logos is known, then all time and existence is known in that measure.<sup>2</sup>

In terms of these conjectures, the importance of psychology for the spectatorial function may be examined in two ways. Many things exist—among them, men. In one cosmological perspective, men may be viewed with utter detachment. They are no more and no less than any other items of the total contents of time and existence. Psychology's reports concerning the nature of these existents, men, are merely reports concerning a set of things in a far more comprehensive set, vaguely called Nature. Any scientist engaged upon inquiry must proceed with functional detachment from the fact that he himself

is a man. He may have to consider how this fact conditions his inquiry activities. He may distrust his passions. He may wish to assist his eyes with instruments. But all such considerations are merely incidental to methodical procedure in the process of inquiry. The philosopher, however, attempting the spectatorial function, must in one perspective upon existence view all things with universal detachment and neutrality. Psychology reports upon something in time and existence. Other sciences make their several reports concerning the parts of time and existence that interest them. Psychology, then, is in this perspective merely one science amid the sciences, and so far there appears no reason for making much ado about it.

Does the exercise of the spectatorial or cosmological function require an additional perspective, a perspective in which psychology is seen to be pre-eminently significant for philosophy? Let us assume that this *may* be the case. After assuming this possibility, we may then ask under what conditions *would* it *be* the case? To indicate these conditions, let us revert to notions found in an earlier stage of our intellectual tradition. I refer to the notions enshrined in such phrases as 'the hierarchy of being,' the 'order of existence from the lower to the higher,' 'the scale of creatures.' No extended consideration of the background of meaning for these phrases need be undertaken. Roughly, they convey the notion that some things are more real than others. Some things possess a larger measure of being. Higher beings represent, distantly, yet more fully than 'lower' things, the plenitude of existence. Things towards the upper levels of the hierarchy represent the actualization of a higher, a richer, set of potentialities. For the sake of brevity, let us call the general standpoint represented in these phrases

<sup>2</sup> In this discussion, I am omitting all considerations of a theological order.

the hierarchical view. Today this view may not seem acceptable, especially in the sciences. It is notable, however, that the hierarchical distinctions seem to be replaced by not dissimilar distinctions. The differences represented in the hierarchical view have been re-formulated rather than abandoned. On the one hand, it may be held that things that exist, molecules, stones, plants, animals, whatever they are, participate equally in existence. With respect to the bare fact that they exist, one thing may be regarded as no more and no less 'real' than another. On the other hand, we observe that many distinctions are made. Higher things are contrasted with lower things. The distinction is applied to chemical substances, to forms of life, to societies and civilizations. The more highly evolved things are distinguished from the less highly evolved. Things are placed in arrangements of degree. The more complex is distinguished from the less complex. The more usual is contrasted with the relatively unusual—and not always in a merely numerical sense. Differences in degree of organization are variously described. The layman concludes, sometimes, that the more highly organized is understood to come about by a gathering into itself of the less highly organized. The former appears to depend upon the prior existence of the latter, or at least logically implies the latter. Insofar as science deals with particulars or singulars, it appears that the treatment of the singular depends upon both the scientific objective at hand and the nature of the singular. For one purpose individual differences are neglected. Even human beings may be just counted. Owing to circumstances, individual differences may elude observation and reckoning. One molecule of water is taken as if identical with every other. It may be the case, one conjectures, that the less complex and less

highly organized may be treated as if individual differences were not significant, with less danger of misrepresentation, than would obtain with respect to the more complex and organized. Heterogeneity may be abstracted from and the heterogeneous reduced to a type of homogeneity, more easily in the case of one kind of thing than in the case of others. Broadly speaking, individuality may be of less consequence, or be treated as if it were of less consequence, with respect to inanimate things than with respect to animate things, and with respect to 'lower' animate things than with respect to 'higher' ones. The unity of the higher thing appears to involve more, and more diverse, factors than that of the lower thing. These considerations and others suggest that the hierarchical view of existence has been translated into new terms.

If we may suppose that scientific consensus would declare man to be the most highly organized, the most complex, the most evolved, the most exceptional natural thing, then this is enough to make men peculiarly significant to the spectator of all time and existence. Can more than this be said? Under what conditions would psychology be indispensably and supremely important for the spectatorial function of philosophy?

To state these conditions conjecturally, let us summarily describe the distinctions indicated above by the simple phrase, the higher and the lower. The question may then be expressed in this fashion: Is the higher a richer source of possible knowledge concerning all time and existence than is the lower? If we assume this to be true, we could then treat the following as promising hypotheses. The higher thing discloses the nature of Nature—the Logos—more amply than does the lower. It has greater evidential value than the lower. It is more fertile in yielding inferences

concerning the Logos of time and existence. It furnishes a sounder basis for speculation. This may be clarified by a supposition. Suppose that man is a higher thing than a stone, and suppose further that, in some sense we cannot presently define, our knowledge of man is as adequate to the constitution of man as our knowledge of (say) granite is adequate to the constitution of granite. If, then, the general principle be sound, the philosopher should be able to penetrate more deeply into the mysteries of all time and existence by exploiting knowledge concerning man than by exploiting knowledge concerning stones. In these conditions, the pre-eminence of psychology is incontestable. If there should be things higher than man, say angels; if these beings were as adequately known as man, then angelology would be the pre-eminent science for the uses of philosophy in its spectatorial function. There is a science of psychology, if not an angelology. In accord with our suppositions, then, this science reveals to the philosophic eye, in principle at least, the extraordinary resources of Nature. Its Logos is made manifest in highest degree by psychological science. Psychology, subject to the provisos stated, is in principle the most philosophically informing of the descriptive sciences.

I have great difficulty in understanding why these conjectures should be discountenanced. From the spectatorial point of view, man, together with all that would not be there if man did not exist, are contents of time and existence and manifestations of its Logos. If man be the central object of psychological investigation, it is man in his complete organization, man concretely and completely viewed in respect of all the factors defining him as a 'higher' being, that psychology seeks to understand. Other anthropocentric sciences are comparatively abstract. Their findings may

be more assured than those of psychology. But this may be due to the fact that they consider man in limited abstract contexts, with little regard for the intricate constitution of this thing. Psychology alone is in principle the complete science of man. Now man is peculiarly the inquiring animal. He achieves sciences and philosophy. He, beyond all other known things, is the *making* animal. He produces algebraic symbols, airplanes, symphonies, governments, laws and moral codes, statues of gods, and Heaven alone knows what else. He exploits the availability of physical nature for incorporating ideas. If there is to be a science of man commensurate with the nature of the thing, all of his capacities for all of his works must be brought into the reckoning. If this is not the task of psychology,—pre-eminently its task,—then to what science is it to be assigned?

The conjectures advanced converge upon this statement: If Nature comprises man, if the 'higher' expresses the internal constitution and resources of Nature more adequately than the 'lower,' then psychology must propose to give knowledge of them more amply, more profoundly, more adequately, and more concretely than any other descriptive science. The very existence of man may be cosmically precarious. Granite could survive more environmental changes than could the race of men. Human beings appear to be confined to a minute region of the cosmos. Because of the temperature of Sirius, no Beethovens and no Leonardos dwell there. Comparatively speaking, there are few human beings and astoundingly many helium particles in the universe. Man, then, is exceptional in that Nature does not produce many men. This, however, is of scant importance. Man is exceptional in extraordinarily significant ways. It is the problem of psychology to make intelligible the constitution of

this exceptional being, and thereby to make intelligible whatever depends on this constitution.

Is it unreasonable to urge that the higher and more exceptional thing has greater significance with respect to the philosopher's effort to discern the Logos of all time and existence? If it be agreed that this is not unreasonable, if man be a very 'high' and exceptional being, then, I submit, the reports of the pre-eminent anthropocentric science have values transcending those of any other science. The Logos of Nature, so far as it is there to be discerned by mind, should be made more intelligibly manifest in man than in the helium particle.<sup>2</sup> It may be far more difficult to secure knowledge of the former than of the latter. Despite this, however, psychology remains for the spectatorial function the supreme natural science, in principle if not yet in accomplishment.

These conjectures<sup>2</sup> may arouse ob-

<sup>2</sup> I cheerfully concede that difficult questions are hidden by my conjectures. For example, if man be described as more intricately constituted and more highly organized than something else, say crystals or molecules, just what does that imply concerning our ability to know these things? And concerning the nature of our knowledge of them? What is involved when we view a thing, say a man, as composed of organized things (say molecules), when in turn the molecules are declared to be more highly organized than the entities of which the molecule is composed? Does the higher organization of the more highly organized thing leave the less highly organized unaffected as to its organization when the less highly organized becomes an integrated 'part' of the former? Is there a sense in which the more complex and highly organized is really simpler, more truly one, than the less complex and organized? Aristotle might urge that there is a difference between what is simpler in itself and what is simpler with respect to our cognitive powers, or what is simpler in an earlier stage of a science and what is simpler in a later stage. In short, what do the distinctions cited amount to in terms of our knowledge, procedures of inquiry, and adequacy of knowledge? Does the high-low distinction mean something like

jections. It may be urged that a science such as physics provides fundamental, basic, knowledge, that physics would provide knowledge more fundamental than that of psychology even were the latter science to achieve the consolidated systematic character and dependability of physics. Perhaps this is the case. But one wonders why it should be the case. In what sense is any body of knowledge more fundamental than another? In what sense is knowledge given by physics more revealing than that given by entomology, or sociology, or psychology? If it has this special value, with respect to what has it this value? Is physics a deeper, more penetrating form of knowledge? Into what, then, does it penetrate more deeply? Many difficulties lurk beneath these questions, and I must necessarily deal with them hurriedly. If in physics there is knowledge in an authentic sense, then knowledge from other sources must accord with that of physics. By definition, however, truth must accord with truth. If there be a truth in psychology, and a truth in physics, and if these are *truths*, then assuredly *each* truth demands that the other be in harmony with it. At any rate, the requirement is that these truths shall not be in discord. One truth no more legitimately voices the demand for accord than does another truth. In this respect, then, how can the truth of physics be more fundamental than a truth of psychology? So far no reason appears for in-

this:—that a stone, say, can be made intelligible by employing fewer principles, by employing principles more easily discovered, than is the case with a 'higher' thing, say a living thing? Is it easier to exhaust, for human knowledge and human purposes, the constitution of the mineral than of man? Circumstances prevent me from exploring the implications of my conjectures. Accordingly, I point to these questions in a footnote to indicate why I describe the discussion as conjectural.



sisting that psychology's findings must be adjusted to the results of physics, rather than the reverse.

For the sake of discussion let us concede that physics has attained some truths, truths final and established in the sense that all further claims alleged to be truths of physics must be consistent with the former. Further, let us assume that the equivalence of matter and energy, however formulated, is such a governing truth. If this truth be called fundamental, in some eulogistic sense, on what grounds should it be so described? It may be called fundamental because of its comprehensive scope. If all time and existence involves matter, then this truth, that matter and energy are equivalent, pertains to all time and existence. On the other hand, if only a portion of all time and existence involves matter, then this truth enlightens us concerning the constitutive character of this portion, but presumably of only this portion. It is a part of the Logos of this portion of Nature, and would somehow be comprised within the Logos of all Nature. Thus the equivalence of matter and energy is a truth about all things in time and existence, or else about a great many things. In either case we should be jubilant. This truth represents a remarkable achievement of the animal that psychology studies.

Maintaining the supposition, we may return to philosophy. In relation to the spectatorial function of philosophy, is this truth of physics fundamental? Is it of commanding importance? Whether this truth holds with respect to all time and existence, or with respect only to a region, what would make it more fundamental for the spectatorial effort? Is this truth more commanding than any truth that psychology possesses or could possess? If the equivalence of matter and energy, or any other truth of physics, penetrated more deeply

into the mystery of existence than any truth of psychology, then assuredly the former would be more significant than the latter. If the truth of physics is more profound than that of psychology, the greater value of the former would be conceded. It would have a privileged position in relation to the spectatorial function. Is it necessary, however, for the philosopher to assume that the truth of physics is the more penetrating, the more profound, whatever these eulogistic terms may mean? I must confess that I see no compelling reason for adopting this position. Just because the assumed truth may reveal something about the Logos of all time and existence, or of a great portion of it, it is not clear that this extensive pertinence gives it cosmological pre-eminence. It seems possible that a truth could be a truth about everything existing, and yet be a relatively superficial truth. It is conceivable that little of Nature's constitution and hidden resources is revealed by a truth about everything existing.

Such a truth may be of great practical importance, in relation to human activities. Man exploits materiality's availability as supplying means to ends. The equivalence of matter and energy, which we have assumed to be an established truth, would more or less remotely condition every effort to control and use physical existence. But this fact may not be of profound significance in the perspective appropriate to philosophy's spectatorial function. If that function be legitimate, its objective is to apprehend, so far as man can accomplish this, the intrinsic constitution of all time and existence. In this sense, then, I am far from persuaded that the equivalence of matter and energy, or any truth of physics, is 'fundamental' or profound or functionally of overwhelming significance. It may represent the fact that men have hit upon meth-

ods whereby existential truths may be secured. But so would truths achieved by bacteriology or psychology. Suppose it were urged that in physics we have some immutable truths, or at least hypotheses of vast scope, remarkable dependability, and predictive fertility. Suppose it were further urged that in some other science, such as psychology, there is no comparable fund of ideas and warranted claims. If we conceded this, what is thereby demonstrated? It is not demonstrated that the results of physics are more profound and philosophically significant. The conceded fact may signify only this, that it is more difficult to obtain such results when inquiring into one region of time and existence than when inquiring into some other region. Or it may mean that truths about all Nature or large portions of it are exceedingly abstract, easier to obtain, and easier to comprehend, than truths reflecting Nature's rich concreteness and its resources for producing the exceptional and highly organized. The facts may warrant us in saying this: It would be astonishing if by now physicists had not arrived at some truths or impressively reliable findings; but it would be even more astonishing if psychologists had by now arrived at *any* truths or *any* dependable findings! The psychologist deals with an entity having an uniquely exceptional status. The truth-finding physicists are themselves instances. Doubtless, even with respect to these instances, insofar as men are physical, the equivalence of matter and energy obtains. Does this enlighten us very much concerning a Nature that in the fullness of time displays its resources by producing physicists, and then by producing psychologists to wonder about them? It is psychology that must enlighten us concerning a Nature which produces a thing that inquires and

makes, a thing that inquires even concerning the Logos of Nature.

With this we are led to consider a different function of philosophy, and so to a new consideration of the relations between philosophy and psychology. Obviously we are referring to the epistemological tradition. But let us avoid the term, and again take advantage of Plato's rhetoric. Man seeks truth, it will be conceded. Perhaps he has never found any truths concerning time and existence. However, he certainly does arrive at opinions concerning them. Plato seems to believe that he may arrive even at *right* opinions. At any rate, Plato distinguishes between right and wrong opinions, and right opinions are said to be the mean between ignorance and knowledge. Man, then, is a producer of opinions and perhaps of right opinions, even if he is not a thing that knows the truth. He is also a maker, but since making somehow depends on opinions, we may confine attention to the production of opinions. He produces opinions concerning all time and existence—perhaps none of these have the dignity of being right—and opinions concerning some of the contents of all time and existence—and some of these may be right. It appears to be a function of the philosopher to consider the production of opinions and the nature of the product. Indeed, perhaps this function is exercised in order to arrive at right opinions concerning all opinions, both right and wrong, concerning such distinctions as those between ignorance and knowledge, and between right and wrong opinions. The philosopher is then in function a critic, an assayer, a reckoner of accounts. Then let us call this task of philosophy its critical function. The philosopher must be the critic of all opinions concerning all time and existence.

Philosophy in this function must become anthropocentric, or must at least

depend upon the results of the pre-eminently anthropocentric science. The process of opinion making must be made intelligible, because the product reflects the conditions of production. Circularity cannot be evaded. To deal effectively with his responsibilities, the philosopher is dependent upon the availability of right opinions concerning the nature of man. Psychology, above all other sciences, is responsible for providing the right opinions required for the exercise of the critical function. Unless philosophy, however, can establish means for discriminating right from wrong opinions, how can either the philosopher or the psychologist determine what psychological results have the dignity of being right opinions? It is not my present purpose to explore this procedural difficulty. My purpose is rather to emphasize its significance for determining the inter-relations of philosophy and psychology. The philosopher must examine critically all bodies of right and wrong opinions. The opinions of psychology, as it were, intervene between the critical function and all other sets of opinions. The task of appraising the opinions constituting physics, say, or economics, biology, sociology, is conditioned by the state of our knowledge concerning the thing that inquires. The ability of the philosopher to arrive at right opinions concerning the values of scientific opinions depends, immediately or remotely, upon psychology's success in finding right opinions concerning man.

The activity of inquiring, and perhaps even the activity of making, may be conducted *as if* the inquirer were not a man. The limitations imposed by the nature of man may be disregarded. This is justifiable as a matter of procedure. Technical devices are employed to implement this procedural disregard. In the end, however, the anthropocentric conditioning of all in-

quiry activity must be reckoned with in the total enterprise of philosophy. The critical function of philosophy tends to engulf its spectatorial function, while in return, the effort to view all time and existence must embrace even the critical function and its dependence upon right opinions concerning the nature of man. No science or art is achieved by a pure intellect functioning serenely outside of all time and existence. Inquiring and making are done by things found within time and existence. If, then, psychology has long possessed an impressive, organized and integrated system of dependable right opinions, the backward state of philosophy is perplexing. It would suggest that philosophers had wilfully neglected to take advantage of the very science upon which their critical function depends. On the other hand, if psychology so far has not achieved a system of right opinions, philosophical failure is to that extent understandable. But in this case the apparent success of other sciences, and psychology's backward state, is somewhat mystifying. Why should psychology be less advanced than some other sciences?

Perhaps some light may be thrown upon this matter by asking what service philosophical thought may render psychology. It may be urged that, so far, psychology has been shown to be indispensable for philosophy, but not philosophy for psychology. If philosophy is peculiarly contributive to psychology, why is this the case? How can the dependence of psychology upon philosophical thought be demonstrated? If preceding conjectures are thought to be reasonable, they suggest answers to these questions. In sum, the central problem of psychology is far more difficult than the central problems of other sciences. Far more is at stake in psychology, not only for human weal or woe but also for knowledge of all time

and existence, than is the case with such sciences as physics, chemistry or bacteriology. To speak figuratively, Nature made manifest its extraordinary complexities and hidden resources, when it produced man, far more completely than when it produced other things. By producing man it produced societies, states, sciences, arts, and philosophies. It placed man before man as an object of inquiry. It thus presented him with a problem of baffling complexity. To satisfy the passion for knowledge, the problem must be investigated. For the sake of human welfare, the nature of man must be made known. The adequacy or inadequacy of psychology's hypotheses is of the utmost significance. Psychology dares not hypostatize abstractions, and represent Nature's most extraordinary accomplishment by caricatures. If one function of philosophy be that of criticizing opinions, then psychology needs philosophical criticism as does no other science.

The problems associated with the older doctrine concerning the hierarchy of existence have been transformed rather than cast aside. In modern form these problems have to do with the relations of the sciences to one another. A curious situation obtains. Higher things as well as lower things may fall within the province of one and the same science. Yet this science may proceed as if the distinction did not exist, or else as if it were not of major importance. Thus the distinction between the living and the non-living thing is virtually non-existent in the context of physics. Even for chemistry it is not of governing significance. It comes to pass, accordingly, that a given science may have to take account of the results of another science, while the latter is not influenced by the former. The biological sciences cannot ignore the conclusions of physics and chemistry,

while the latter can proceed as if the former had nothing to contribute to them. Psychology, especially, must consider the results of many other sciences, for all of them comprise men within their perspectives upon time and existence. The burden of the psychologist is all the more onerous. The other sciences advance opinions that more or less incidentally are opinions concerning man. But none of them is a science, or *the* science, concerning man. Psychology, accordingly, has difficulty in maintaining its integrity. It tends to become dissipated into innumerable inquiries concerning abstracted and isolated details. To a certain degree this may be not only expedient but necessary. The integration of results, however, becomes all the more difficult. Here again, the philosopher may insist, philosophical criticism is seen to be indispensable for the progress of psychology. It is the function of philosophy to speculate about the relations between the various sciences. It must somehow adjust the results of each to the others. It must call attention to the diversities found in all time and existence. It must determine the scope of the sets of opinions constituting the many sciences, and appraise their relative significance. By exercising these critical functions philosophical thought furnishes indispensable support for the psychologist's efforts towards integration.

The psychologist is led, sometimes, to ape other scientists. He even takes over-seriously their ideas just as if these other sciences had the towering importance possessed by his own science. He then needs the emancipating influence of both the spectatorial and critical functions of philosophy. His proposal, to make intelligible the most exceptional thing in all time and existence, is one of extreme audacity. The psychologist all the more needs the sobering counsels of philosophy. He may



retort to this by saying that the quests of philosophy are even more audacious! This may be granted. It is one indication of the importance of psychology for philosophy. And there is a rejoinder. Little is expected of philosophy, and especially because it is so dependent upon psychology's search for right opin-

ions about that thing which may perhaps provide the best clues concerning the mystery of all time and existence. Philosophy may wait. But philosophers cannot. Philosophers, accordingly, must urge psychologists to greater efforts,—efforts marked at once by humility and audacity.

1

This monograph summarizes the attitudes of the average German civilian. It tells about his wartime experiences, his hopes and fears, his loyalties and hates. Identification with the Nazi party had more to do with morale than any of the other influences studied.

Psychological Monograph No. 276    \$1.25

**OBSERVATIONS** *on the*  
*Characteristics and Distribution*  
*of* **GERMAN NAZIS**

by  
**Helen Peak**  
Connecticut College

**American Psychological Association**  
1515 Massachusetts Avenue N. W.  
Washington 5, D. C.

# PSYCHOLOGICAL REVIEW

YEAR	VOLUME	AVAILABLE NUMBERS						PRICE PER NUMBER	PRICE PER VOLUME
1894	1	-	2	-	4	5	6	\$1.00	\$4.00
1895	2	-	2	3	4	5	6	\$1.00	\$5.00
1896	3	1	-	-	-	-	-	\$1.00	\$1.00
1897	4	1	-	-	-	5	6	\$1.00	\$3.00
1898	5	1	2	3	4	5	6	\$1.00	\$5.50
1899	6	1	-	3	-	5	6	\$1.00	\$4.00
1900	7	1	-	3	4	5	-	\$1.00	\$4.00
1901	8	1	2	-	-	5	-	\$1.00	\$3.00
1902	9	-	2	-	4	-	-	\$1.00	\$2.00
1903	10	1	2	3	4	5	6	\$1.00	\$5.50
1904	11	1	-	3	4	5	6	\$1.00	\$5.00
1905	12	1	2	3	4	5	6	\$1.00	\$5.50
1906	13	-	-	3	4	5	6	\$1.00	\$4.00
1907	14	1	2	3	-	5	6	\$1.00	\$5.00
1908	15	1	-	3	-	-	-	\$1.00	\$2.00
1909	16	1	2	3	4	5	6	\$1.00	\$5.50
1910	17	1	2	3	4	5	6	\$1.00	\$5.50
1911	18	1	2	3	4	5	6	\$1.00	\$5.50
1912	19	1	2	3	4	5	6	\$1.00	\$5.50
1913	20	1	2	3	4	5	6	\$1.00	\$5.50
1914	21	-	2	3	4	5	6	\$1.00	\$5.00
1915	22	1	2	3	4	5	6	\$1.00	\$5.50
1916	23	1	2	-	4	5	6	\$1.00	\$5.00
1917	24	1	2	-	4	5	6	\$1.00	\$5.00
1918	25	-	2	3	4	5	6	\$1.00	\$5.00
1919	26	1	2	3	4	5	6	\$1.00	\$5.50
1920	27	1	2	3	4	5	6	\$1.00	\$5.50
1921	28	-	2	3	4	5	6	\$1.00	\$5.00
1922	29	1	-	-	4	-	-	\$1.00	\$2.00
1923	30	1	2	3	4	5	6	\$1.00	\$5.50
1924	31	1	2	3	4	5	6	\$1.00	\$5.50
1925	32	1	2	3	-	5	-	\$1.00	\$4.00
1926	33	1	2	3	4	5	6	\$1.00	\$5.50
1927	34	1	2	3	4	5	6	\$1.00	\$5.50
1928	35	1	2	3	4	5	6	\$1.00	\$5.50
1929	36	1	2	3	4	5	6	\$1.00	\$5.50
1930	37	1	2	3	4	5	6	\$1.00	\$5.50
1931	38	1	2	3	4	5	6	\$1.00	\$5.50
1932	39	1	2	3	4	5	6	\$1.00	\$5.50
1933	40	1	2	3	4	5	6	\$1.00	\$5.50
1934	41	1	2	3	4	5	6	\$1.00	\$5.50
1935	42	1	2	3	4	5	6	\$1.00	\$5.50
1936	43	1	2	3	4	5	6	\$1.00	\$5.50
1937	44	1	2	3	4	5	6	\$1.00	\$5.50
1938	45	1	2	3	4	5	6	\$1.00	\$5.50
1939	46	1	2	3	4	5	6	\$1.00	\$5.50
1940	47	1	2	3	4	5	6	\$1.00	\$5.50
1941	48	1	2	3	4	5	6	\$1.00	\$5.50
1942	49	1	2	3	4	5	6	\$1.00	\$5.50
1943	50	1	2	3	4	5	6	\$1.00	\$5.50
1944	51	1	2	3	4	5	6	\$1.00	\$5.50
1945	52	1	2	3	4	5	6	\$1.00	\$5.50
1946	53	1	2	3	4	5	6	\$1.00	\$5.50
1947	54	1	2	3	4	5	6	\$1.00	\$5.50
1948	55	By Subscription, \$5.50						\$1.00	

List price, Volumes 1 through 54

\$265.50

30% Discount

79.65

Net price, Volumes 1 through 54

\$185.85

Information about the Psychological Review: from 1894 to 1908 many numbers are out of print, as shown in the table. After 1908, only ten numbers are out of print. After 1920, no numbers are out of print.

The journal has been published with six numbers a year throughout its history.

Information about prices: the Psychological Review has the uniform price of \$5.50 per volume and \$1.00 per issue. For incomplete volumes, the price is \$1.00 for each available number. For foreign postage, \$2.25 per volume should be added. The American Psychological Association gives the following discounts on orders for any one journal:

10% on orders of \$ 50.00 and over

20% on orders of \$100.00 and over

30% on orders of \$150.00 and over

Current subscriptions and orders for back numbers should be addressed to:

**AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.**

1515 Massachusetts Avenue, N. W.

Washington 5, D. C.



